

```

ID Q9JHLO PRELIMINARY; PRT; 203 AA.
AC Q9JHLO;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE WSCR15 protein (WSCR15).
GN WSCR15 OR WSCR15.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Doyle J.L., Desliva U., Miller W., Green E.;
RT "Divergent Human and Mouse Orthologs of a Novel Gene (WSCR15/Wscr15)
RT Reside within the Genomic Interval Commonly Deleted in Williams
RT Syndrome."
RL Cytogenet. Cell Genet. 0:0-0(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ES-129/SVJ;
RA Martindale D.W., Wilson M.D., Wang D., Burke R.D., Chen X.;
RA Durando V., Koop B.F.;
RT "Comparative genomic sequence analysis of the Williams syndrome region
RT (LIMK1-RFC2) of human chromosome 7q11.23."
RL Mamm. Genome 0:0-0(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RA Green E.D.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF257136; AAF9153.1; -
DR EMBL; AF139987; AAF75558.1; -
DR EMBL; AF289664; AAF9331.1; -
DR MGD; MGI:1926479; Wscr15.
SQ SEQUENCE 203 AA; 22876 MW; 1B21A87D8FBAE097 CRC64;

Query Match 33.3%; Score 5; DB 11; Length 203;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STPES 12
DB 148 STPES 152

RESULT 197
Q20551 PRELIMINARY; PRT; 204 AA.
AC Q20551;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Hypothetical 23.0 kDa protein.
GN F48B9.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peleoderinae; Caenorhabditis.
OX NCBI_TaxId=6339;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium."
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Miller N.;
RT "The sequence of C. elegans cosmid F48B9."

```

```

RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission."
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; U40955; AAA81747.2; -.
DR HSSP; P26367; 6PAX.
DR Interpro; IPR001523; Paired_box.
DR Pfam; PF00292; PAX; 1.
DR SMART; SM00351; PAX; 1.
KW Hypothetical protein.
SQ SEQUENCE 204 AA; 22990 MW; 40C2B6B96A8A0F8C CRC64;

Query Match 33.3%; Score 5; DB 5; Length 204;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STPES 12
DB 9 STPES 13

RESULT 198
Q9N7O9 PRELIMINARY; PRT; 206 AA.
AC Q9N7O9;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Possible hypothetical protein mj1404 (fragment).
GN Lm28.132.
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxId=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FRIDLIN;
RA Murphy L., Quail M., Harris D., Rajandream M., Ivens A., Barrell B.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL390935; CAC00878.1; -.
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 206 AA; 23695 MW; 33A5F4B26A12A0F0 CRC64;

Query Match 33.3%; Score 5; DB 5; Length 206;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STPES 12
DB 146 STPES 150

RESULT 199
Q35872 PRELIMINARY; PRT; 206 AA.
AC Q35872;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE ABC-type heme transporter subunit.
GN ORF206.
OS Lycopersicon esculentum (Tomato).
OC Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxId=4081;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Mailli R., Marziani A.,
RA Miltischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shin P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Tortum M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana.";
RT Nature 408:816-820(2000).
DR EMBL: AC068901; AAG50893.1; -.
DR InterPro: IPR004252; Transposase_24.
DR Pfam: PF03004; Transposase_24; 1.
KW Hypothetical protein.
SQ SEQUENCE 201 AA; 23130 MW; 15C49205EFFBC614 CRC64;

Query Match
Best Local Similarity 33.3%; Score 5; DB 10; Length 201;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPHR 7
DB 100 LGPHR 104

RESULT 193
O8UY82 PRELIMINARY; PRT; 201 AA.
AC O8UY82;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE PVIT.
GN L2.
OS Simian adenovirus 25.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=175567;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PAN 9;
RX MEDLINE=21548291; PubMed=11689642;
RA Farina S.F., Gao G.P., Xiang Z.Q., Rux J.J., Burnett R.M.,
RA Alvira M.R., Marsh J., Ertl H.C., Wilson J.M.;
RT "Replication-defective vector based on a chimpanzee adenovirus.";
RN J. Virol. 75:11603-11613(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=PAN 9;
RA Alvira M.R.;
RT Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF394196; AAI35522.1; -.
DR InterPro: IPR004912; Adeno_VIT.
DR Pfam: PF03228; Adeno_VIT; 1.
SQ SEQUENCE 201 AA; 22292 MW; B1F7BDBEF8CF47DF CRC64;

Query Match
Best Local Similarity 33.3%; Score 5; DB 12; Length 201;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 HRSTP 10
DB 109 HRSTP 113

RESULT 194
O8UI48 PRELIMINARY; PRT; 201 AA.
AC O8UI48;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical protein Atu0452.

```

```

GN ATU0452.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176293;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D., St.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58.";
RT Science 294:2317-2323(2001).
DR EMBL: AE009015; AAL41471.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 201 AA; 21160 MW; 141FD5F09C263762 CRC64;

Query Match
Best Local Similarity 33.3%; Score 5; DB 16; Length 201;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRAA 15
DB 163 ESRAA 167

RESULT 195
O8I611 PRELIMINARY; PRT; 203 AA.
AC O8I611;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Beta-galactosidase (Fragment).
OS Carica papaya (Papaya).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustosids II; Brassicales; Caricaceae; Carica.
OX NCBI_TaxID=3649;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SOLO;
RA D'Innocenzo M., Nasclmento J.R.O., LaJoie F.M.;
RT "Purification and Characterization of Carica papaya fruit beta-
RT galactosidase.";
RT Submitted (JUL-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF079874; AAC28739.1; -.
DR InterPro: IPR001944; GH_35.
DR Pfam: PF01301; Glyco_hydro_35; 1.
DR PRINTS: PR00742; GLHYDRASE35.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 203 AA; 23243 MW; 970778B7174E90D4 CRC64;

Query Match
Best Local Similarity 33.3%; Score 5; DB 10; Length 203;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPE 11
DB 2 RSTPE 6

RESULT 196
O9JHLO

```



```

OX NCBI_TaxID=109760;
RN [1]
RP SEQUENCE FROM N.A.
RA Forger L., Ustlova J., Wang Z., Huss V.A.R., Lang F.B.F.;
RT "Hydrophidulum curvatum: a linear mitochondrial genome, tRNA editing,
RL Mol. Biol. Evol. 0:0-0(2001).";
DR EMBL: AF404303; AAK84247.1; -.
SQ SEQUENCE 196 AA; 21719 MW; E7942C53C0522006 CRC64;

Query Match
Best Local Similarity 100.0%; Score 5; DB 8; Length 196;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPES 12
DB 172 STPES 176

RESULT 189
O9UB9 PRELIMINARY; PRT; 197 AA.
AC O9UB9;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE D36266.1.1 (Continues in Em:AL031681 as d3138B7.3) (Fragment).
GN D3138B7.3
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Laird G.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL031681; CAB3959.1; -.
FR NON_TER 197
SQ SEQUENCE 197 AA; 21361 MW; 44A6F2D9CC3FEB1F CRC64;

Query Match
Best Local Similarity 100.0%; Score 5; DB 4; Length 197;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPES 12
DB 185 STPES 189

RESULT 190
O9Y974 PRELIMINARY; PRT; 197 AA.
ID O9Y974
AC O9Y974;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein APE2411.
GN APE2411.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococciales;
OC Desulfurococcaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=K1.
RC MEDLINE=99310339; PubMed=10382966;
RA Kawaiibayashi Y., Hino Y., Horikawa H., Yamazaki S., Hatake Y.,
RA Jii-no K., Takahashi M., Sekine M., Baba S.-I., Anka A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;

```

```

RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
RW EMBL: AP000064; BAB1426.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 197 AA; 20339 MW; DADDEB1214F83806 CRC64;

Query Match
Best Local Similarity 100.0%; Score 5; DB 17; Length 197;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLP 5
DB 17 SHLP 21

RESULT 191
O9H6L6 PRELIMINARY; PRT; 201 AA.
ID O9H6L6
AC O9H6L6;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE CDNA: FLJ22151 fis, clone HRC0111.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
RA Suzuki Y., Ohashi M., Nishi T., Shibahara T., Tanaka T.,
RA Nakamura Y., Isono T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK025804; BAB15240.1; -.
SQ SEQUENCE 201 AA; 20931 MW; 44CED5BFC28333 CRC64;

Query Match
Best Local Similarity 100.0%; Score 5; DB 4; Length 201;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPES 12
DB 60 STPES 64

RESULT 192
O9C7M6 PRELIMINARY; PRT; 201 AA.
ID O9C7M6
AC O9C7M6;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Hypothetical 23.1 kDa protein.
GN F103.5.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=CV. COLUMBIA;
RC MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altieri H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Etlg P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizlar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,

```

```

RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA da Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Trufi D., Tsai S.M., Tsunako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
DR EMBL: AF004019; AAF84809.1; -
DR InterPro: IPR000415; Nitroreductase.
DR Pfam: PF00881; Nitroreductase; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 191 AA; 21582 MW; D9F6B5466B149EE CRC64;

Query Match 33.3%; Score 5; DB 16; Length 191;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HLGPH 6
DB 154 HLGPH 158

RESULT 186
O95C95 PRELIMINARY; PRT; 192 AA.
AC O95C95:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Small ribosomal protein 4 (Fragment).
GN RPS4.
OS Huperzia lucidula (shining club moss).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Lycopodiophyta; Lycopodiales; Lycopodiaceae; Huperzia.
OX NCBI_TaxID=37429;
RN [1]
RP SEQUENCE FROM N.A.
RA Fryer K.M., Schneider H., Smith A.R., Cranfill R., Wolf P.G.,
RA Hunt J.S., Sipes S.D.;
RT "Horsetails and ferns are a monophyletic group and the closest living
RT relatives to seed plants.";
RL Nature 0:0-0(2001).
DR EMBL: AF313605; AAU26206.1; -
DR InterPro: IPR001912; Ribosomal_S4.
DR Pfam: PF00163; Ribosomal_S4; 1.
DR Pfam: PF01479; S4; 1.
DR TIGRPFAMs: TIGR01017; rpsd_bact; 1.
DR PROSITE: PS00632; RIBOSOMAL_S4; UNKNOWN_1.
KW Chloroplast; Ribosomal protein.
FT NON_TER 1
SQ SEQUENCE 192 AA; 22312 MW; 683CFEE60130679E CRC64;

Query Match 33.3%; Score 5; DB 8; Length 192;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRA 14
DB 133 PESRA 137

RESULT 187
O9L113 PRELIMINARY; PRT; 195 AA.
AC O9L113:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

```

```

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE GblAAFA3227.1 (ATg18560/K24M9.5).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-COLIMBIA;
RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL Submitted (Mar-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-COLIMBIA;
RX MEDLINE=20363099; PubMed=10907853;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the regions of 4,251,695 bp covered by ninety p1,
RT TAC and BAC clones.";
RL DNA Res. 7:217-221(2000).
RN [3]
RP SEQUENCE FROM N.A.
RA Kim C.J., Chen H., Cheuk R., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Tortum M., Yamada A.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis ORF clones.";
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Kim C.J., Chen H., Cheuk R., Koesema E., Meyers M.C., Shin P.,
RA Banh J., Bowser L., Carninci P., Chung M.K., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Tortum M., Yamada K.,
RA Yu G., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis cDNA clones.";
RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP001303; BAB02214.1; -
DR EMBL: AY054138; AAU6799.1; -
DR EMBL: AF380644; AAK55725.1; -
DR InterPro: IPR000923; BlueCu_1.
DR PROSITE: PS00196; COPPER_BLUE.
SQ SEQUENCE 195 AA; 21876 MW; 854707849DADCAD1 CRC64;

Query Match 33.3%; Score 5; DB 10; Length 195;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
DB 3 RSTPE 7

RESULT 188
O95OR1 PRELIMINARY; PRT; 196 AA.
AC O95OR1:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Orf196.
GN ORF196.
OS Spizellomyces punctatus.
OS Mitochondrion.
OC Eukaryota; Fungi; Chytridiomycota; Spizellomycetales;
OC Spizellomycetaceae; Spizellomyces.

```

RC STRAIN-0157.H7 / RIND 0509952;
 RX MEDLINE-21156231; PubMed-11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,
 Kubera S., Shiba T., Hattori M., Shinagawa H.,
 RT *Complete genome sequence of enterohemorrhagic Escherichia coli
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22(2001).
 DR EMBL: AE006457; AAK16936.1; -
 DR EMBL: AP002556; BAB35238.1; -
 KW Complete proteome.
 SQ SEQUENCE 189 AA; 21388 MW; EE5461D5021DAE54 CRC64;

Query Match 33.3%; Score 5; DB 16; Length 189;
 Best Local Similarity 100.0%; Pred. No. 4.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPES 12
 11111
 DB 24 STPES 28

RESULT 183
 P70368 PRELIMINARY; PRT; 190 AA.
 AC P70368;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE SEBOX.
 GN OG9X.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C;
 RX MEDLINE-97008065; PubMed-8855241;
 RA Rovescalli A.C., Asob S., Nirenberg M.;
 RT "Cloning and characterization of four murine homeobox genes.";
 RT Proc. Natl. Acad. Sci. U.S.A. 93:10691-10696(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C;
 RX MEDLINE-20381309; PubMed-10922053;
 RA Cinguranta M., Rovescalli A.C., Kozak C.A., Nirenberg M.;
 RT "Mouse Sebox homeobox gene expression in skin, brain, oocytes, and
 RT two-cell embryos.";
 RT Proc. Natl. Acad. Sci. U.S.A. 97:8904-8909(2000).
 DR EMBL: U65068; AAC52829.2; -
 DR HSSP: P06601; 1FTL.
 DR TRANSFAC: T03309; -
 DR MGD: MGI:108012; OG9X.
 DR InterPro: IPR001356; Homeobox.
 DR Pfam: PF00046; homeobox.1.
 DR ProDom: PD000010; Homeobox.1.
 DR SMART: SM00389; HOX.1.
 DR PROSITE: PS00027; HOMEBOX_1; UNKNOWN_1.
 DR PROSITE: PS50071; HOMEBOX_2; 1.
 SQ SEQUENCE 190 AA; 20391 MW; 4D473850504129A3 CRC64;

Query Match 33.3%; Score 5; DB 11; Length 190;
 Best Local Similarity 100.0%; Pred. No. 4.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHR 7
 11111
 DB 15 LGPHR 19

RESULT 184

Q9J329 PRELIMINARY; PRT; 191 AA.
 AC Q9J329;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE Wbscr5 alternative spliced product (Williams-Beuren syndrome
 DE chromosome region 5 homolog) (human).
 GN WBSR5.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ES-129/SV11;
 RA Martindale D.W., Wilson M.D., Wang D., Burke R.D., Chen X.,
 RA Duranilo V., Koop B.F.;
 RT "Comparative genomic sequence analysis of the Williams syndrome region
 RT (11M1-RFC2) of human chromosome 7q11.23.";
 RL Mamm. Genome 0:0-0(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RT Submitted (Apr-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF139987; AAF75559.1; -
 DR EMBL: BC005804; AA05804.1; -
 DR MGD: MGI:1926479; Wbscr5.
 SQ SEQUENCE 191 AA; 21408 MW; 1B874CA4D55A01E6 CRC64;

Query Match 33.3%; Score 5; DB 11; Length 191;
 Best Local Similarity 100.0%; Pred. No. 5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPES 12
 11111
 DB 136 STPES 140

RESULT 185
 Q9PBX8 PRELIMINARY; PRT; 191 AA.
 AC Q9PBX8;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hypothetical protein Xf2007.
 GN Xf2007.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 OC Xylella.
 OX NCBI_Taxid=2371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-9A5C;
 RX MEDLINE-20365717; PubMed-10910347;
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvaranga R., Alves L.M.C., Araya J.E., Baia G.S., Battista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Brites M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carer H.,
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
 RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Fetto J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,
 RA Garlier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hohnsels J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.B., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhani A. Jr., Nodrega F.G., Nunes L.R., Oliveira M.A.,

```

DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE SEBOX.
GN SEBOX.
OS Rattus norvegicus (Rat)..
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY;
RA MEDLINE=20381309; PubMed=10922053;
RA Cinqunta M., Rovescalli A.C., Kozak C.A., Nirenberg M.;
RT "Mouse Sbox homeobox gene expression in skin, brain, oocytes, and
RT two-cell embryos."
RL Proc. Natl. Acad. Sci. U.S.A. 97:8904-8909(2000).
DR EMBL; AF284338; AAG14459.1; -.
DR HSSP; P06601; 1FJL.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR Prodom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; UNKNOWN_1.
DR PROSITE; PSS0071; HOMEBOX_2; 1.
FT VARIANT 149 149 W->C.
FT VARIANT 153 153 E->G.
SQ SEQUENCE 188 AA; 20196 MW; C7F649EBC35F9B0 CRC64;

Query Match 33.3%; Score 5; DB 11; Length 188;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPHR 7
Db 15 LGPHR 19

RESULT 181
Q9ZUE1 PRELIMINARY; PRT; 189 AA.
AC Q9ZUE1;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DE 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE Hypothetical 20.9 kDa protein F508.9 (F28C11.16).
GN F508.9.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA Vysotskaya V.S., Schwartz J.R., Yu G., Toriumi M., Liu S., Lenz C.,
RA Li J., Kremenetskaia I., Luros J., Altafi H., Gonzalez A., Araujo R.,
RA Buehler E., Conn L., Conway A.B., Dunn P., Hansen N., Huizar L.,
RA Kim C., Palm C.J., Rowley D., Shinn P., Walker M., Davis R.W.,
RA Ecker J.R., Federspiel N.A., Theologis A.;
RT "Arabidopsis thaliana chromosome 1 BAC F508 sequence."
RT Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA Theologis A.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [1]
RP SEQUENCE FROM N.A.
RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,
RA Kim C., Altafi H., Bel O., Chin C., Chlou J., Choi E., Conn L.,
RA Conway A., Gonzales A., Hansen N., Howling B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharzky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,

```

```

RA Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC F28C11 from chromosome
RT 1."
RN Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [1]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
RA Shinn P., Altafi H., Bel B., Chin C., Chlou J., Choi E., Conn L.,
RA Conway A., Gonzalez A., Hansen N., Howling B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu J., Liu S., Mukharzky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N., Theologis A.,
RA Ecker J.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA Khan S., Kim C., Altafi H., Bel B., Chin C., Chlou J., Choi E.,
RA Conn L., Conway A., Gonzalez A., Hansen N., Howling B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharzky N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC005990; AAC98009.1; -.
DR EMBL; AC007945; AAF79589.1; -.
KW Hypothetical protein.
SQ SEQUENCE 189 AA; 20864 MW; D5450D695471748D CRC64;

Query Match 33.3%; Score 5; DB 10; Length 189;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRRA 15
Db 68 ESRRA 72

RESULT 182
Q8XAL7 PRELIMINARY; PRT; 189 AA.
AC Q8XAL7;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Orf, hypothetical protein.
GN 26020 OR ECS1815.
OS Escherichia coli O157:H7.
OC Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / EDL933 / ATCC 700927;
RA MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G., III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postal G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobleck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RT Nature 409:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.

```

RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-UTERUS;
 RA Strausberg R.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC004262; AA04262.1; -.
 FT NON_TER
 SQ SEQUENCE 183 AA; 22169 MW; 39646D10E25B960F CRC64;

Query Match
 Best Local Similarity 33.3%; Score 5; DB 4; Length 183;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
 DB 114 RSTPE 118

RESULT 177
 ID 082013 PRELIMINARY; PRT: 183 AA.
 AC 082013;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Hypothetical protein Alr0113.
 GN Alr0113.
 OS Anabaena sp. (strain PCC 7120).
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 OX NCBI_TaxId=103690;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21595285; PubMed=11759840;
 RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
 RA Matsumoto M., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kohara M., Matsumoto M., Muraki A.,
 RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
 RA Yasuda M., Tabata S.;
 RT "Complete genomic sequence of the filamentous nitrogen-fixing
 cyanobacterium Anabaena sp. strain PCC 7120.";
 RL DNA Res. 8:205-213(2001).
 DR EMBL; AP003581; BAB77637.1; -.
 DR InterPro: IPR000182; GCM5Acetyltransf.
 DR Pfam: PF00583; Acetyltransf. 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 183 AA; 20939 MW; 02AD828D376F3397 CRC64;

Query Match
 Best Local Similarity 33.3%; Score 5; DB 16; Length 183;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPES 12
 DB 9 STPES 13

RESULT 178
 ID 081624 PRELIMINARY; PRT: 185 AA.
 AC 081624;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE F8M12.19 protein.
 GN F8M12.19.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxId=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;

RA Washu;
 RT "The A. thaliana Genome Sequencing Project.";
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Madsen C., Graves T., Cotton M., Modde T.;
 RT "The sequence of A. thaliana F8M12.";
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Waterston R.;
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF080118; AAC33959.1; -.
 SQ SEQUENCE 185 AA; 22148 MW; BC826D37F18DE04B CRC64;

Query Match
 Best Local Similarity 33.3%; Score 5; DB 10; Length 185;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHR 7
 DB 126 LGPHR 130

RESULT 179
 ID 09KY24 PRELIMINARY; PRT: 187 AA.
 AC 09KY24;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Putative lipoprotein.
 GN SC07535 OR SC8612.11.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
 OX NCBI_TaxId=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kieser T., Larke L., Oliver K., O'Neill S.,
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrall B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 coelicolor A3(2).";
 RL Nature 417:141-147(2002).
 DR EMBL; AL355753; CAB90893.1; -.
 KW Lipoprotein.
 SQ SEQUENCE 187 AA; 19027 MW; ECDF24D053649944 CRC64;

Query Match
 Best Local Similarity 33.3%; Score 5; DB 16; Length 187;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRRA 15
 DB 102 ESRRA 106

RESULT 180
 ID 09ERS8 PRELIMINARY; PRT: 188 AA.
 AC 09ERS8;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)

RA Tissot C., Nissen J., Wechti N.;
 RT "Molecular cloning of a new interferon-inductible PMV nuclear bodies-
 associated protein."
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Pentecost B.T.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE-B-CELL;
 RA Strausberg R.;
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: X89773; CA61915.2; -;
 DR EMBL: U88964; AAB53416.1; -;
 DR EMBL: BC007922; AA07922.1; -;
 DR InterPro: IPR000520; Exonuclease.
 DR Pfam: PF00929; Exonuclease; 1.
 DR SMART: SM00479; EXOIII; 1.
 SO SEQUENCE 181 AA; 20363 MW; 24519CB52CEA581 CRC64;

Query Match 33.3%; Score 5; DB 4; Length 181;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHR 7
 |||||
 DB 17 LGPHR 21

RESULT 174
 O92LB0 PRELIMINARY; PRT; 181 AA.
 ID O92LB0;
 AC O92LB0;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Hypothetical protein R03164.
 GN R03164 OR SMC03781.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 NC NCB1_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RA Capela D., Barloy-Hubier F., Gouzy J., Bothe G., Ampe F., Batut J.,
 RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
 RA Pohl T., Portetle D., Puhler A., Purnelle B., Ransperger U.,
 RA Renard C., Thebaud P., Vandenbol M., Weidner S., Gallbert F.;
 RT "Analysis of the chromosome sequence of the legume symbiont
 Sinorhizobium meliloti strain 1021."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
 DR EMBL: AL591793; CAC47743.1; -;
 DR InterPro: IPR003742; DUF163.
 DR Pfam: PF02590; DUF163; 1.
 KW Hypothetical protein; Complete proteome.
 SO SEQUENCE 181 AA; 19513 MW; F45BDD60303R75 CRC64;

Query Match 33.3%; Score 5; DB 16; Length 181;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRAA 15
 |||||
 DB 67 ESRAA 71

RESULT 175
 O9LOE7 PRELIMINARY; PRT; 181 AA.
 ID O9LOE7

AC O9LOE7;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical protein SC04695.
 GN SC04695 OR SCD31.20.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 NC NCB1_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Saunders D.C., Harris D.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA MEDLINE=97000351; PubMed=8843436;
 RA Redenbach M., Kieser H.M., Denapate D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome."
 RL Mol. Microbiol. 21:77-96(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Godle A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 coelicolor A3(2)."
 RL Nature 417:141-147(2002).
 DR EMBL: AL161803; CAB82063.1; -;
 DR InterPro: IPR001932; PP2C-like.
 DR InterPro: IPR001230; Prey1-site.
 DR SMART: SM00331; PP2C_S1G; 1.
 DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.
 KW Hypothetical protein.
 SO SEQUENCE 181 AA; 18977 MW; B915019A97814814 CRC64;

Query Match 33.3%; Score 5; DB 16; Length 181;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRAA 15
 |||||
 DB 124 ESRAA 128

RESULT 176
 O9BTA6 PRELIMINARY; PRT; 183 AA.
 ID O9BTA6;
 AC O9BTA6;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE Similar to cactin (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NC NCB1_TaxID=9606;

RP SEQUENCE FROM N.A.
 RC STRAIN-ZOBELL ATCC 14405;
 RX MEDLINE=97107629; PubMed=6950369;
 RA Glockner A.B., Zumft W.G.;
 RT "Sequence and analysis of an internal 9.7-kb segment from the 30-kb
 RT denatification gene cluster of *Pseudomonas stutzeri*.";
 RL Biochim. Biophys. Acta 1277:6-12(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ZOBELL ATCC 14405;
 RX MEDLINE=96028114; PubMed=7588711;
 RA Palmedo G., Seither P., Koerner H., Matthews J.C., Burkhalter R.S.,
 RA Tinkovych R., Zumft W.G.;
 RT "Resolution of the nrd locus for heme d1 synthesis of cytochrome cdi
 RT (respiratory nitrite reductase) from *Pseudomonas stutzeri*.";
 RL Eur. J. Biochem. 232:737-746(1995).
 DR EMBL: 273914; CA98149.1; -;
 DR InterPro: IPR000298; CytC_oxdse_III.
 DR Pfam: PF00510; COX3; 1.
 DR ProDom: PD000382; CytC_oxdse_III; 1.
 DR PROSITE: PS50253; COX3; 1.
 SQ SEQUENCE 175 AA; 19485 MW; 24E2E3430ACD7ABF CRC64;

Query Match 33.3%; Score 5; DB 2; Length 175;
 Best Local Similarity 100.0%; Pred. No. 4.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 11 ESRAA 15
 |||||
 Db 32 ESRAA 36

RESULT 171
 Q9VKY6 PRELIMINARY; PRT; 176 AA.
 AC Q9VKY6;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE CG13141 protein.
 GN CG13141.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards R., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Crawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferlita S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Gload A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Jaisli M., Kaulush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Jaisli M., Kaulush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003628; AAF52920.1; -;
 DR FlyBase: FBgn0032226; CG13141.
 DR InterPro: IPR003654; Homeo_OAR.
 SQ SEQUENCE 176 AA; 17810 MW; 568496C33A4DF25F CRC64;

Query Match 33.3%; Score 5; DB 5; Length 176;
 Best Local Similarity 100.0%; Pred. No. 4.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 7 RSTPE 11
 |||||
 Db 3 RSTPE 7

RESULT 172
 O80ZT0 PRELIMINARY; PRT; 176 AA.
 AC O80ZT0;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DE Similar to unknown (Protein for IMAGE:3508182).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-EYE;
 RA Strausberg R.;
 RL Submitted (MAR-2002) to the EMBL/Genbank/DBJ databases.
 DR EMBL: BC024802; AAH24802.1; -;
 SQ SEQUENCE 176 AA; 19102 MW; 9B309FA9E585AC6 CRC64;

Query Match 33.3%; Score 5; DB 11; Length 176;
 Best Local Similarity 100.0%; Pred. No. 4.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 10 PESRA 14
 |||||
 Db 114 PESRA 118

RESULT 173
 Q00441 PRELIMINARY; PRT; 181 AA.
 AC Q00441; 000586;
 DT 01-JUL-1997 (TREMBLrel. 04, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE ISG20 protein (Interferon stimulated gene) (20KD).
 GN ISG20 OR HEM45.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.

RC TISSUE-LUNG;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC014429; AAI14429.1; -;
DR InterPro: IPR000369; ISK_Channel.
DR Pfam: PF02060; ISK_Channel; 1.
SQ SEQUENCE 170 AA; 18439 MW; 4F6C94F87BF71B52 CRC64;
Query Match 33.3%; Score 5; DB 4; Length 170;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 11 ESRAA 15
Db 22 ESRAA 26
RESULT 167
O52386 PRELIMINARY; PRT; 172 AA.
AC O52386;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE RepB homolog.
OS Lactobacillus rhamnosus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Lactobacillaceae; Lactobacillus.
OX NCBI_TaxID=47715;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC7469;
RA Kawai S., Shimamura K., Kashiuchi N., Yamamoto K., Hideniko K.;
RT "Lactobacillus rhamnosus surface located protein."
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF037091; AAB91419.1; -;
SQ SEQUENCE 172 AA; 19966 MW; F8B392A3FB9E2FEF4 CRC64;
Query Match 33.3%; Score 5; DB 2; Length 172;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 8 STRES 12
Db 156 STRES 160
RESULT 168
O9RR06 PRELIMINARY; PRT; 173 AA.
AC O9RR06;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical protein DR2432.
GN DR2432.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RL;
RA MEDLINE=20036896; PubMed=1056726;
RX White O., Eissen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwin M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus radiodurans RL.";

RL Science 286:1571-1577(1999).
DR EMBL: AF002073; AAF11973.1; -;
DR TIGR: DR2432; -;
DR InterPro: IPR000182; GCN5acetyltransf.
DR Pfam: PF00583; Acetyltransf; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 173 AA; 18755 MW; 42F8A6DFA4E4A77D CRC64;
Query Match 33.3%; Score 5; DB 16; Length 173;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 11 ESRAA 15
Db 80 ESRAA 84
RESULT 169
O9W6F2 PRELIMINARY; PRT; 174 AA.
AC O9W6F2;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Protein A1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99190706; PubMed=10090728;
RA Lee R.M., Gilet G., Burnside J., Thomas S.J., Neiman P.;
RT "Role of Nr13 in regulation of programmed cell death in the bursa of
RL Fabricius."
RL Genes Dev. 13:718-728(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Sofer L., Burnside J.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF120211; AAD31645.1; -;
DR InterPro: IPR000712; BCL2_BH.
DR InterPro: IPR002475; BCL2_family.
DR Pfam: PF00452; Bcl-2; 1.
DR SMART: SM00337; BCL; 1.
DR PROSITE: PS5062; BCL2_FAMILY; 1.
DR PROSITE: PS01258; BH2; 1.
SQ SEQUENCE 174 AA; 20095 MW; 4880F463DB22B352 CRC64;
Query Match 33.3%; Score 5; DB 13; Length 174;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHLGP 5
Db 24 SHLGP 28
RESULT 170
P95548 PRELIMINARY; PRT; 175 AA.
AC P95548;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE ORF175 protein.
GN ORF175.
OS Pseudomonas stutzeri (Pseudomonas perfectomarina).
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=316;
RN [1]


```

RESULT 163
006236 PRELIMINARY; PRT: 168 AA.
AC 006236;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Similar to SDHAP.
CN YEL164W OR L9632.1.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=5288C;
RX MEDLINE=97313267; PubMed=9169871;
RA Johnston M., Hillier L., Riles L., Albermann K., Andre B., Ansorge W.,
RA Benes V., Bruckner M., Delius H., Dubois E., Dusterhoft A.,
RA Entian K.D., Floeth M., Goffeau A., Hebling U., Heumann K.,
RA Heuss-Nelzel D., Hilbert H., Hilger F., Kiehl K., Kotter P.,
RA Louis E.U., Messenguy F., Mewes H.W., Miosga T., Mostl D.,
RA Muller-Auer S., Neutwich U., Obermaier B., Piravandi E., Pohl T.M.,
RA Portetelle D., Purnelle B., Reichmann S., Rieger M., Rinke M., Rose M.,
RA Scharfe M., Scherens B., Scholler P., Schwager C., Schwarz S.,
RA Underwood A.P., Urrestarazu L.A., Vandenbol M., Verhasselt P.,
RA Vierendeels F., Voet M., Volckaert G., Voss H., Wambolt R., Wedler E.,
RA Wedler H., Zimmermann F.K., Zollner A., Hani J., Hohnselt J.D.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
RL Nature 387:0-0(0).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=5288C;
RA Vaudin M.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=5288C;
RA Waterston R.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=5288C;
RA Cherry J.M.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U51921; AAB67488.1; -.
DR SGD: S0004154; YIR164W.
SQ SEQUENCE 168 AA; 18632 MW; 296AE82172DE9DCD CRC64;

Query Match 33.3%; Score 5; DB 3; Length 168;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRA 14
DB 51 PESRA 55

RESULT 164
097194 PRELIMINARY; PRT: 168 AA.
AC 097194;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Hypothetical 19.1 kDa protein.
CN L2385.02.
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=FRIEDLIN;
RA Ivens A.C., Murphy L., Quail M., Harris D., Oliver K., Lawson D.,
RA Rajandream M.A., Barrell B.G.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FRIEDLIN;
RX MEDLINE=98146435; PubMed=9477341;
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA Smith D.F.;
RT "A physical map of the Leishmania major Friedlin genome.";
RL Genome Res. 8:135-145(1998).
DR EMBL: AL139794; CAC22659.1; -.
KW Hypothetical protein.
SQ SEQUENCE 168 AA; 19126 MW; 950E3FAE9715EA98 CRC64;

Query Match 33.3%; Score 5; DB 5; Length 168;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRAA 15
DB 131 ESRAA 135

RESULT 165
08WMG9 PRELIMINARY; PRT: 170 AA.
AC 08WMG9;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Potassium voltage-gated channel-like protein.
CN KCND4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY, AND HEART;
RA Hul R., Jeng S., Lin C., Ma L., Zhen Y.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY065987; AAL49979.1; -.
DR InterPro: IPR000369; ISK_Channel.
DR PRINTS: PR00168; KCNECHANNEL.
SQ SEQUENCE 170 AA; 18397 MW; 1C6FBCF87298F6C0 CRC64;

Query Match 33.3%; Score 5; DB 4; Length 170;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRAA 15
DB 22 ESRAA 26

RESULT 166
096CC4 PRELIMINARY; PRT: 170 AA.
AC 096CC4;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Similar to potassium voltage-gated channel, Isk-related subfamily,
DE gene 4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

```

```

DT 01-MAR-2001 (TREMBlrel. 16, last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, last annotation update)
DE Vng0121h.
GN VNG0121h.
OS Halobacterium sp. (strain NRC-1).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahatras G.G., Bergquist B., Pan M.,
RA Shukla H.D., Laszky S.R., Baliga N.S., Thorsson V., Shroga J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithausen B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Beck R.F., Pohlischer M., Spudich J.L., Jung K.-H.,
RA Alam M., Freltas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.,
RT "Genome sequence of Halobacterium species NRC-1."
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL: AE004979; MAG18745.1;
KW Complete proteome.
SQ SEQUENCE 163 AA; 17127 MW; 8283FB54AC56F530 CRC64;

Query Match
Best Local Similarity 33.3%; Score 5; DB 17; Length 163;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STPES 12
DB 29 STPES 33

RESULT 160
024121 PRELIMINARY: PRT; 165 AA.
AC 024121;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, last annotation update)
DE Transcription factor.
OS Nicotiana glauca (N. glauca) (Leadwort-leaved tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4092;
RN [1]
RP SEQUENCE FROM N.A.
RA Borisjuk N.V.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: Y09106; CAI70323.1;
DR InterPro: IPR002715; NAC.
DR Pfam: PF01849; NAC; 1.
SQ SEQUENCE 165 AA; 17859 MW; 04DECAC5C15B38BD CRC64;

Query Match
Best Local Similarity 33.3%; Score 5; DB 10; Length 165;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PSRA 14
DB 120 PSRA 124

RESULT 161
09CCZ0 PRELIMINARY: PRT; 165 AA.
AC 09CCZ0;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)
DE Conserved hypothetical protein.

```

```

GN ML2654.
OS Mycobacterium leprae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churche C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holtroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus."
RL Nature 409:1007-1011(2001).
DR EMBL: AL583926; CAC32186.1;
DR Leptoma; ML2654;
DR InterPro: IPR004844; S/T-phosphatase.
DR Pfam: PF00149; Metallophos; 1.
KW Complete proteome.
SQ SEQUENCE 165 AA; 18036 MW; 4DC7C826CBE6D833 CRC64;

Query Match
Best Local Similarity 33.3%; Score 5; DB 16; Length 165;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRA 15
DB 48 ESRA 52

RESULT 162
09LH78 PRELIMINARY: PRT; 167 AA.
AC 09LH78;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
DE Similarity to transposon protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RX MEDLINE=20363099; PubMed=10907853;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the regions of 4,251,695 bp covered by ninety P1,
RT TAC and BAC clones."
RL DNA Res. 7:217-221(2000).
DR EMBL: AP002061; BAB02643.1;
DR EMBL: AP002061; BAB02643.1;
SQ SEQUENCE 167 AA; 19719 MW; FF4B30BA17333ADE CRC64;

Query Match
Best Local Similarity 33.3%; Score 5; DB 10; Length 167;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STPES 12
DB 5 STPES 9

```

Db 84 TPESTR 88

RESULT 155

058905

ID 058905; PRELIMINARY; PRT; 160 AA.

AC 058905; 01-AUG-1998 (TREMBLrel. 07, Created)

DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE Hypothetical protein PH1191.

GN PH1191.

OS Pyrococcus horikoshii.

OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;

OC Pyrococcus.

OX NCBI_TaxID=53953;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=OT3;

RX MEDLINE=98344137; PubMed=9679194;

RA Kawarabayashi Y., Sawada M., Horikawa H., Hino Y.,

RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,

RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,

RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushiida N., Oguchi A.,

RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,

RA Masuchi Y., Shizuya H., Kikuchi H.;

RT "Complete sequence and gene organization of the genome of a hyper-

RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";

RL DNA Res. 5:55-76(1998).

DR EMBL; AP000005; BAA30291.1; -

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 160 AA; 18718 MW; A21DD7F21CD0E95 CRC64;

Query Match Best Local Similarity 33.3%; Score 5; DB 17; Length 160; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 HRSTP 10

Db 4 HRSTP 8

RESULT 156

09N1L3

ID 09N1L3; PRELIMINARY; PRT; 161 AA.

AC 09N1L3; 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE Olfactory receptor (Fragment).

GN EF0145.

OS Eulemur fulvus (brown lemur).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Strepsirhini; Lemnidae; Eulemur.

OX NCBI_TaxID=15151;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20183981; PubMed=10706615;

RA Rouquier S., Blancher A., Giorgi D.;

RT "The olfactory receptor gene repertoire in primates and mouse:

RT Evidence for reduction of the functional fraction in primates.";

RL Proc. Natl. Acad. Sci. U.S.A. 97:2870-2874(2000).

DR EMBL; AF197711; AAF40354.1; -

DR InterPro: IPR000276; GPCR_Rhodpsn.

DR Pfam: PF00001; 7tm_1; 1.

DR PROSITE: PS50262; G_PROTEIN_RECPT_FL_2; 1.

KW Receptor.

FT NON_TER 1

FT NON_TER 161

SQ SEQUENCE 161 AA; 17687 MW; 3A67FBF76F3B069 CRC64;

Query Match Best Local Similarity 33.3%; Score 5; DB 6; Length 161; Matches 100.0%; Pred. No. 4.3e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRAA 15

Db 148 ESRAA 152

RESULT 157

053000

ID 053000; PRELIMINARY; PRT; 163 AA.

AC 053000;

DT 01-JUN-1998 (TREMBLrel. 06, Created)

DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)

DT 01-AUG-1998 (TREMBLrel. 07, Last annotation update)

DE CshD protein.

GN CSHD.

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Escherichia.

OX NCBI_TaxID=562;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=PE;

RA Clark C.A., Manning P.A.;

RL Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.

DR EMBL; X71971; CAA50788.1; -

SQ SEQUENCE 163 AA; 17562 MW; FA7A7ADFELAF649C CRC64;

Query Match Best Local Similarity 33.3%; Score 5; DB 2; Length 163; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRAA 15

Db 75 ESRAA 79

RESULT 158

096129

ID 096129; PRELIMINARY; PRT; 163 AA.

AC 096129;

DT 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE Unknown (Protein for IMAGE:4301250) (Fragment).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=UTERUS;

RA Strausberg R.;

RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC007864; AAH07864.1; -

DR InterPro: IPR004822; Histone_core.

FT NON_TER 1

FT NON_TER 1

SQ SEQUENCE 163 AA; 18449 MW; 70AD52C6D9575A47 CRC64;

Query Match Best Local Similarity 33.3%; Score 5; DB 4; Length 163; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STPEs 12

Db 5 STPEs 9

RESULT 159

09HS08

ID 09HS08; PRELIMINARY; PRT; 163 AA.

AC 09HS08;

DT 01-MAR-2001 (TREMBLrel. 16, Created)

GN RSBW.
OS Staphylococcus epidermidis.
OC Bacteria: Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1457;
RX MEDLINE=21172880; PubMed=11274123;
RA Knobloch J.K.-M., Bartscht K., Sabottke A., Rohde H., Feucht H.H.,
Mack D.;
RT "Biotin Formation by Staphylococcus epidermidis Depends on Functional
RsbU, an Activator of the sigB Operon: Differential Activation
RT Mechanisms Due to Ethanol and Salt Stress.";
RL J. Bacteriol. 183:2624-2633(2001).
DR EMBL: AF274004; AAC23812.1; -
DR InterPro: IPR003594; ATPbind_ATPase.
DR InterPro: IPR004359; HIS_KIN_sig.
DR Pfam: PF02518; HATPase_c; 1.
DR SMART: SM00387; HATPase_c; 1.
SQ SEQUENCE 159 AA; 17922 MW; EA6A49C76218B78F CRC64;

Query Match 33.3%; Score 5; DB 2; Length 159;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGP 5
DB 97 SHLGP 101

RESULT 152
OBSVS5 PRELIMINARY; PRT; 159 AA.
ID 08VS5;
AC 08VS5;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE RsbW.
GN RSBW.
OS Staphylococcus epidermidis.
OC Bacteria: Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21562662; PubMed=11705980;
RA Kies S., Otto M., Vuong C., Gotsz F.;
RT "Identification of the sigB Operon in Staphylococcus epidermidis:
RT Construction and Characterization of a sigB Deletion Mutant.";
RL Infect. Immun. 69:7933-7936(2001).
DR EMBL: AF359562; AL37942.1; -
DR InterPro: IPR003594; ATPbind_ATPase.
DR InterPro: IPR004359; HIS_KIN_sig.
DR Pfam: PF02518; HATPase_c; 1.
DR PROSITE: PS50109; HIS_KIN; 1.
SQ SEQUENCE 159 AA; 17924 MW; 06B481BAA60CF34 CRC64;

Query Match 33.3%; Score 5; DB 2; Length 159;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGP 5
DB 97 SHLGP 101

RESULT 153
ID 004218 PRELIMINARY; PRT; 159 AA.
AC 004218;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Putative G-protein-coupled receptor (Fragment).
GN GCRL
OS Brassica napus (Rape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3708;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98189351; PubMed=9512416;
RA Plakidou-Dymock S., Dymock D., Hookey R.;
RT "A higher plant seven-transmembrane receptor that influences
RT sensitivity to cytokinins.";
RL Curr. Biol. 8:315-324(1998).
DR EMBL: U95144; AAC49963.1; -
DR InterPro: IPR000832; GPCR_secretin.
DR PROSITE: PS50261; G_PROTEIN_RECEP_F2_4; 1.
DR Receptor.
KW NON_TER
FT NON_TER 1 159
SQ SEQUENCE 159 AA; 18568 MW; DBBBAED0C02DBF50 CRC64;

Query Match 33.3%; Score 5; DB 10; Length 159;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGP 5
DB 61 SHLGP 65

RESULT 154
ID 09F484 PRELIMINARY; PRT; 160 AA.
AC 09F484;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE FhuA (Fragment).
GN FhuA.
OS Alteromonas sp. (strain O-7).
OC Bacteria: Proteobacteria; gamma subdivision; Alteromonadaceae;
OC Alteromonas.
OX NCBI_TaxID=29458;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20422186; PubMed=10966390;
RA Tsujibo H., Miyamoto K., Okamoto T., Orihoshi H., Inamori Y.;
RT "A Serine Protease-Encoding Gene (aprtII) of Alteromonas sp. Strain O-7
RT Is Regulated by the Iron Uptake Regulator (Fur) Protein.";
RL Appl. Environ. Microbiol. 66:3778-3783(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O-7;
RX MEDLINE=94124007; PubMed=8294010;
RA Tsujibo H., Miyamoto K., Tanaka K., Kawai M., Tainaka K., Imada C.,
Okami Y., Inamori Y.;
RT "Cloning and sequence of an alkaline serine protease-encoding gene
RT from the marine bacterium Alteromonas sp. strain O-7.";
RL Gene 136:247-251(1993).
DR EMBL: AB040412; BAB13364.1; -
DR HSP: P06971; 2FCF.
FT NON_TER 160 160
SQ SEQUENCE 160 AA; 17578 MW; 95A66CCFE671DCDA CRC64;

Query Match 33.3%; Score 5; DB 2; Length 160;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESR 13
DB 11111

```

FT  NON_TER      1      1
SQ  NON_TER      154     154
    SEQUENCE      154 AA; 16954 MW; 252998E71528BDFD CRC64;

Query Match
Best Local Similarity 100.0%; Score 5; DB 12; Length 154;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  7 RSTPE 11
    11111
Db  142 RSTPE 146

RESULT 148
Q48591 PRELIMINARY; PRT; 155 AA.
ID  Q48591;
AC  Q48591;
DT  01-NOV-1996 (TREMBlrel. 01, Created)
DT  01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE  N of 16S rRNA gene (5'end) (Fragment).
OS  Lactococcus lactis.
OC  Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
CX  Streptococcaceae; Lactococcus.
NX  NCBI_TaxID=1358;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=NCDO 712;
RX  MEDLINE=94065631; PubMed=7504067;
RA  Beresford T., Condon S.;
RT  "Physiological and genetic regulation of rRNA synthesis in
    Lactococcus.";
RL  J. Gen. Microbiol. 139:2009-2017(1993).
DR  EMBL; X65713; CAA46629.1; -.
FT  NON_TER      1      1
SQ  SEQUENCE      155 AA; 18651 MW; 2CFABE9A5FF7ABF CRC64;

Query Match
Best Local Similarity 100.0%; Score 5; DB 2; Length 155;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  8 STPE 12
    11111
Db  97 STPE 101

RESULT 149
Q9M9A2 PRELIMINARY; PRT; 156 AA.
ID  Q9M9A2;
AC  Q9M9A2;
DT  01-OCT-2000 (TREMBlrel. 15, Created)
DT  01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE  01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE  F27J15.21.
OS  Arabidopsis thaliana (Mouse-ear cress).
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC  eustosids II; Brassicales; Brassicaceae; Arabidopsids.
NX  NCBI_TaxID=3702;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  Shin P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,
    Kim C., Altafi H., Bel O., Chin C., Chiu J., Choi E., Conn L.,
    Conway A., Gonzales A., Hansen N., Howling B., Koo T., Lam B., Lee J.,
    Lenz C., Li J., Liu A., Liu K., Liu S., Mukharzky N., Nguyen M.,
    Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
    Torlund M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
    Ecker J.R.;
RT  "Genomic sequence for Arabidopsis thaliana BAC F27J15 from chromosome
    I";
RL  Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AC016041; AAF69704.1; -.
SQ  SEQUENCE      156 AA; 17611 MW; 8192BBBD5D7829E CRC64;

```

```

Query Match
Best Local Similarity 100.0%; Score 5; DB 10; Length 156;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  7 RSTPE 11
    11111
Db  3 RSTPE 7

RESULT 150
Q8UE18 PRELIMINARY; PRT; 157 AA.
ID  Q8UE18;
AC  Q8UE18;
DT  01-JUN-2002 (TREMBlrel. 21, Created)
DT  01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DE  01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE  Nitrogen regulatory protein PII.
GN  GLNB OR ATU1769 OR AGR_C_3352.
OS  Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC  Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
CX  Rhizobiaceae; Rhizobium.
NX  NCBI_TaxID=176299;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  MEDLINE=21608550; PubMed=11743193;
RA  Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA  Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,
RA  Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boyce D., Sr.,
RA  Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA  Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA  Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,
RA  Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA  Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA  Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA  Nester E.W.;
RT  "The genome of the natural genetic engineer Agrobacterium tumefaciens
    C58.";
RL  Science 294:2317-2323(2001).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  MEDLINE=21608551; PubMed=11743194;
RA  Goodner B., Hinkle G., Gattung S., Muller N., Blanchard M.,
RA  Gourollo B., Goldman B.S., Cao Y., Askenzi M., Halling C., Mullin L.,
RA  Houmiel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA  Wollam C., Allinger M., Doughly D., Scott C., Lappas C., Markelz B.,
RA  Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,
RA  Clelo C., Slater S.;
RT  "Genome sequence of the plant pathogen and biotechnology agent
    Agrobacterium tumefaciens C58.";
RL  Science 294:2323-2328(2001).
DR  EMBL; AE009133; AAI42768.1; ALT_INIT.
DR  EMBL; AE008098; AAK87538.1; -.
KW  Complete Proteome.
SQ  SEQUENCE      157 AA; 17464 MW; F52E8AD60A7EE93 CRC64;

Query Match
Best Local Similarity 100.0%; Score 5; DB 16; Length 157;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  10 PESRA 14
    11111
Db  5 PESRA 9

RESULT 151
Q9F7V2 PRELIMINARY; PRT; 159 AA.
ID  Q9F7V2;
AC  Q9F7V2;
DT  01-MAR-2001 (TREMBlrel. 16, Created)
DT  01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE  01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DR  RSDW.
SQ  SEQUENCE      159 AA; 17611 MW; 8192BBBD5D7829E CRC64;

```

RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
 RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
 RA Perry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
 RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
 RA Metcalf W.W., Birren B.,
 RT "The genome of *Methanoscacia acetivorans* reveals extensive metabolic
 RT and physiological diversity." ;
 RL Genome Res. 12:532-542(2002).
 DR EMBL: AE010765; AAM04426.1; -.
 KW Complete proteome.
 SQ SEQUENCE 150 AA; 16485 MW; 51E988774E444EB1 CRC64;

Query Match 33.3%; Score 5; DB 17; Length 150;
 Best Local Similarity 100.0%; Pred. No. 4.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGPFR 5
 DB 98 SHLGP 102

RESULT 144

OBROPS PRELIMINARY; PRT; 152 AA.
 AC Q8ROP5;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (EC
 DE 4.6.1.12).
 GN YGBB.
 OS *Thermus thermophilus*.
 OC Bacteria; *Thermus*/Deinococcus group; Deinococci; Thermales;
 OC Thermaceae; *Thermus*.
 OX NCBI_TaxID=274;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HB8;
 RA Kishida H., Wada T., Unzai S., Kuzuyama T., Terada T., Shiroyu M.,
 RA Yokoyama S., Tame J.R., Park S.,
 RT "2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase from *Thermus*
 RT *thermophilus* HB8." ;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB082126; BAB86885.1; -.
 KW Lyase.
 SQ SEQUENCE 152 AA; 16520 MW; 0114C0E440DC28F3 CRC64;

Query Match 33.3%; Score 5; DB 2; Length 152;
 Best Local Similarity 100.0%; Pred. No. 4.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPFR 7
 DB 105 LGPFR 109

RESULT 145

O9RELI PRELIMINARY; PRT; 154 AA.
 AC Q9RELI;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE RecA protein (Fragment).
 GN RECA.
 OS *Micrococcus lylae*.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Micrococcales; Micrococcaceae; *Micrococcus*.
 OX NCBI_TaxID=1273;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 27566;
 DR MEDLINE=20378653; PubMed=10919806;

RA van Waasbergen L.G., Balkwill D.L., Crocker F.H., Bjornstad B.N.,
 RA Miller R.V.;
 RT "Genetic diversity among *Arthrobacter* species collected across a
 RT heterogeneous series of terrestrial deep-subsurface sediments as
 RT determined on the basis of 16S rRNA and recA gene sequences." ;
 RL Appl. Environ. Microbiol. 66:3454-3463(2000).
 DR EMBL: AF214778; AAF25425.1; -.
 FT NON_TER 1 1
 FT NON_TER 154 154

SQ SEQUENCE 154 AA; 17504 MW; 1B6C50129389DCBE CRC64;

Query Match 33.3%; Score 5; DB 2; Length 154;
 Best Local Similarity 100.0%; Pred. No. 4.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPFR 7
 DB 143 LGPFR 147

RESULT 146

Q47394 PRELIMINARY; PRT; 154 AA.
 AC Q47394;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE ORF1.
 DE *Escherichia coli*.
 OS *Escherichia coli*.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96109443; PubMed=8619599;
 RA Noguchi N., Emura A., Matsuyama H., Ohara K., Sasatsu M., Kono M.,
 RT "Nucleotide sequence and characterization of erythromycin resistance
 RT determinant that encodes macrolide 2'-phosphotransferase I in
 RT *Escherichia coli*." ;
 RL Antimicrob. Agents Chemother. 39:2359-2363(1995).
 DR EMBL: D16251; BAA03774.2; -.
 KW Lyase.
 SQ SEQUENCE 154 AA; 16915 MW; D72E1B90301B6FD9 CRC64;

Query Match 33.3%; Score 5; DB 2; Length 154;
 Best Local Similarity 100.0%; Pred. No. 4.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRA 14
 DB 147 PESRA 151

RESULT 147

Q9DS07 PRELIMINARY; PRT; 154 AA.
 AC Q9DS07;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE Polyprotein (Fragment).
 OS Foot-and-mouth disease virus Asia 1.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Aphovirus.
 OX NCBI_TaxID=110195;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ASIA1/IND 63/72;
 RA Manju G., Venkataramanan R., Gurusurthy C.B., Hemadri D., Tosh C.,
 RA Sanyal A.;
 RT "Nucleotide sequence of 2B gene of foot-and-mouth disease virus
 RT serotype Asia-1." ;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF207524; AAG35703.1; -.

```

DR EMBL; AP003596; BAB76014.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 145 AA; 15732 MW; D16371D9C4EBA237 CRC64;

Query Match
Best Local Similarity 100.0%; Score 5; DB 16; Length 145;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STRES 12
DB 71 STRES 75

RESULT 140
O9MCM1 ID O9MCM1 PRELIMINARY; PRT; 148 AA.
AC O9MCM1.
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE ORF9.
GN ORF9.
OS Streptococcus thermophilus bacteriophage 7201.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=112023;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20088830; PubMed=10620678;
RA Stanley E., Walsh L., van der Zwet A., Fitzgerald G.F.,
van Sinderen D.;
RT "Identification of four loci isolated from two Streptococcus
thermophilus phage genomes responsible for mediating bacteriophage
resistance.";
RL FEWS Microbiol. Lett. 182:271-277(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Stanley E., Walsh L., Fitzgerald G.F., van Sinderen D.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF145054; AAF26608.1; -.
DR HSSP; P02339; 1EXG.
DR InterPro: IPR000424; SSB_protein.
DR Pfam: PF00436; SSB; 1.
DR TIGRfams: TIGR00621; ssb; 1.
SQ SEQUENCE 148 AA; 16669 MW; 0DD3D7C394A53FF9 CRC64;

Query Match
Best Local Similarity 100.0%; Score 5; DB 9; Length 148;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRAA 15
DB 104 ESRAA 108

RESULT 141
O9Y413 ID O9Y413 PRELIMINARY; PRT; 150 AA.
AC O9Y413.
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE Hypothetical 18.0 kDa protein (Fragment).
GN DKFZP566F0546.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tissue-Kidney.
RA Ottenwelder B., Obermaier B., Meyers H.W., Gassenhuber J., Wiemann S.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL050075; CAB43258.1; -.

```

```

KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 150 AA; 18001 MW; 5F4E4DE03153CA65 CRC64;

Query Match
Best Local Similarity 100.0%; Score 5; DB 4; Length 150;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHR 7
DB 36 LGPHR 40

RESULT 142
O923U3 ID O923U3 PRELIMINARY; PRT; 150 AA.
AC O923U3.
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE TRP5 (Fragment).
GN Cavia porcellus (Guinea pig).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Hystriognathi; Cavidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RA Ong H.L., Breton H.M., Harland M.L., Barritt G.J.;
RT "Expression of mRNA encoding Trp cation channels in guinea pig airway
smooth muscle cells.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF395787; AAK84034.1; -.
DR InterPro: IPR002111; Cat_channel_Trp.
DR InterPro: IPR000636; M-channel_nlg.
DR InterPro: IPR001865; Ribosomal_S2.
DR Pfam: PF00520; Ion_trans; 1.
DR PROSITE; PS00962; RIBOSOMAL_S2_1; UNKNOWN_1.
FT NON_TER
SQ SEQUENCE 150 AA; 17068 MW; B719BE32C5804913 CRC64;

Query Match
Best Local Similarity 100.0%; Score 5; DB 11; Length 150;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGP 5
DB 23 SHLGP 27

RESULT 143
O8TS13 ID O8TS13 PRELIMINARY; PRT; 150 AA.
AC O8TS13.
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Predicted protein.
GN MA0994.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
Fitzhugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
Allen N., Naylor J., Stange-Thomann N., Deatellano K., Johnson R.,
Linton L., McMan P., McKernan K., Talamas J., Tirrell A., Ye W.,
Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,

```

DR EMBL: U77596; AAC60341.1; -.
DR InterPro: IPR000663; Natr_peptide.
DR Pfam: PF00212; ANP; 1.
DR PRINTS: PR00710; NATPEPTIDES.
DR SMART: SM00183; NAT_pep; 1.
DR PROSITE: PS00263; NATRIURETIC_PEP; 1.
SQ SEQUENCE 139 AA; 14881 MW; 426287771870E1FB CRC64;

Query Match 33.3%; Score 5; DB 13; Length 139;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRAA 15
DB 127 ESRAA 131

RESULT 136
Q9NM50 PRELIMINARY; PRT; 143 AA.
AC Q9NM50;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE Possible hypothetical 63.6 kDa protein in aqpz-cspd intergenic region
DE (fragment).
GN LM26.242.
OS Leishmania major.
OC Eukaryota; Eukaryozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-FRIEDLIN.
RA Murphy L., Quail M., Harris D., Rajandream M., Ivens A., Barrell B.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL160493; CAB97860.1; -.
FT NON_CER 1
SQ SEQUENCE 143 AA; 15685 MW; 3287611FDF805A6B CRC64;

Query Match 33.3%; Score 5; DB 5; Length 143;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPE 11
DB 44 RSTPE 48

RESULT 137
Q942D1 PRELIMINARY; PRT; 143 AA.
AC Q942D1;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE B1060H01.8 protein (OSUNB0036G09.14 protein).
GN B1060H01.8 OR OSUNB0036G09.14.
OS Oryza sativa (Rice), and
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530, 39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RA "Oryza sativa nippobare(GAS) genomic DNA, chromosome 1, BAC
RT clone:B1060H01.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;

RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nippobare(GAS) genomic DNA, chromosome 1, BAC
RT clone:OSUNB0036G09.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP003560; BAB68075.1; -.
DR EMBL: AF003509; BAB89938.1; -.
SQ SEQUENCE 143 AA; 15266 MW; F25DC4382B14EE0 CRC64;

Query Match 33.3%; Score 5; DB 10; Length 143;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PHRST 9
DB 26 PHRST 30

RESULT 138
Q9Z3G2 PRELIMINARY; PRT; 145 AA.
AC Q9Z3G2;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE CytM.
GN CytM.
OS Synechococcus sp. (strain PCC 7942) (Anacystis nidulans R2).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=1140;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PCC7942;
RA Malakhov M.P., Murata N.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: U39811; AAD00004.1; -.
DR InterPro: IPR00345; CytC_heme_bind.
DR InterPro: IPR003088; Cyt_C1.
DR Pfam: PF00034; Cytochrome_C; 1.
DR PROSITE: PS00190; CYTOCHROME_C; 1.
SQ SEQUENCE 145 AA; 15545 MW; 8D04DE32CA07A730 CRC64;

Query Match 33.3%; Score 5; DB 2; Length 145;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRAA 15
DB 31 ESRAA 35

RESULT 139
O8YP84 PRELIMINARY; PRT; 145 AA.
AC O8YP84;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein A114315.
GN A114315.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).;

RESULT 132

09CXT5 PRELIMINARY: PRT: 138 AA.

AC 09CXT5:

DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

DE 2310047B19Rik protein.

GN 2310047B19Rik.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=EMBRYONIC HEAD;

RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I., Saito T., Okazaki Y., Gojopori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleschmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schriml L.M., Stanbil F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nodone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whitaker C., Wilmink L., Wyshew-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsuki S., Hayashizaki Y.

RT "Functional annotation of a full-length mouse cDNA collection."

RL Nature 409:685-690(2001).

DR EMBL: AK014010; BAB29112.1; MGD: MGI:1914212; 2310047B19Rik.

SO SEQUENCE 138 AA; 15116 MW; 6195918D98ABD236 CRC64;

Query Match 33.3%; Score 5; DB 11; Length 138;

Best Local Similarity 100.0%; Pred. No. 3.8e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGP 5

DB 89 SHLGP 93

RESULT 133

08WYS2 PRELIMINARY: PRT: 139 AA.

AC 08WYS2:

DT 01-MAR-2002 (TREMBLrel. 20, Created)

DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE Hypothetical 15.7 kDa protein.

GN PP9943.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Zhou X.M., Zhang P.P., Jiang H.Q., Huang Y., Qin W.X., Zhao X.T., Wan D.F., Gu J.R.;

RT "Novel human cDNA clones with function of inhibiting cancer cell growth."

RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF318380; AAL55887.1; -.

KW Hypothetical protein.

SQ SEQUENCE 139 AA; 15662 MW; 9DEF6B0103456A0 CRC64;

Query Match 33.3%; Score 5; DB 4; Length 139;

Best Local Similarity 100.0%; Pred. No. 3.8e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 GPHRS 8

DB 123 GPHRS 127

RESULT 134

08S071 PRELIMINARY: PRT: 139 AA.

AC 08S071:

DT 01-JUN-2002 (TREMBLrel. 21, Created)

DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)

DE P0678F11.7 protein.

GN P0678F11.7.

OS Oryza sativa (japonica cultivar-group).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzaceae; Oryza.

OX NCBI_TaxID=39947;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. NIPPONBARE;

RA Sasaki T., Matsumoto T., Yamamoto K.;

RT "Oryza sativa nippobare(GA3) genomic DNA, chromosome 1, PAC clone:P0678F11."

RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AP003437; BAB6100.1; -.

SO SEQUENCE 139 AA; 16397 MW; B656B7692D3341CB CRC64;

Query Match 33.3%; Score 5; DB 10; Length 139;

Best Local Similarity 100.0%; Pred. No. 3.8e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 PHRST 9

DB 37 PHRST 41

RESULT 135

P79799 PRELIMINARY: PRT: 139 AA.

AC P79799:

DT 01-MAY-1997 (TREMBLrel. 03, Created)

DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE Natriuretic peptide.

OS Micrurus corallinus (Brazilian coral snake).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Lepidodactylia; Squamata; Scleroglossa; Serpentes; Colubroidae;

OC Elapidae; Elapinae; Micrurus.

OX NCBI_TaxID=54390;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=VENOM GLAND;

RA Ho P.L., Soares M.B., Yamane T., Raw I.;

RT "Reverse biology applied to Micrurus corallinus, a South American coral snake."

RL J. Toxicol. Toxin. Rev. 14:327-337(1995).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=VENOM GLAND;

RX MEDLINE=98092299; PubMed=9432002;

RA Ho P.L., Soares M.B., Maack T., Gimenez I., Puerto G., Furtado M.F., Raw I.;

RT "Cloning of an unusual natriuretic peptide from the South American coral snake Micrurus corallinus."

RL Eur. J. Biochem. 250:144-149(1997).

PROMOTER (BY SIMILARITY).
 CC EMBL: AY033233; AAK64199.1; -
 DR InterPro: IPR001831; HIV_Tat.
 DR Pfam: PF00539; Tatf_1.
 KM Activator: Nuclear protein; RNA-binding; Transcription regulation.
 SO SEQUENCE 131 AA; 14742 MW; F93ID793B0CF1DBB CRC64;

Query Match
 Best Local Similarity 100.0%; Score 5; DB 15; Length 131;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STEPS 12
 Db 28 STEPS 32

RESULT 129
 ID 08X104 PRELIMINARY; PRT; 134 AA.
 AC 08X104;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hypothetical protein CPE2320.
 GN CPE2320.
 OS Clostridium perfringens.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
 OC Clostridiales; Clostridiaceae; Clostridium.
 OX NCBI_TaxID=1502;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=13 / TYPE A;
 RX PubMed=11792842;
 RA Shimizu T., Ohnari K., Hirakawa H., Onshima K., Yamashita A.,
 RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.,
 RT "Complete genome sequence of Clostridium perfringens, an anaerobic
 RT flesh-eater.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
 DR EMBL: AP003193; BAB82026.1; -
 KM Hypothetical protein; Complete proteome.
 SO SEQUENCE 134 AA; 15780 MW; 13450C3B61886CF4 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 5; DB 16; Length 134;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPFR 7
 Db 43 LGPFR 47

RESULT 130
 ID 052214 PRELIMINARY; PRT; 135 AA.
 AC 052214;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Hydroxylaminobenzene mutase.
 GN HABA.
 OS Pseudomonas pseudocalligenes.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=330;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JS45;
 RX MEDLINE=20336447; PubMed=10877793;
 RA Davis J.K., Paoli G.C., He Z., Nadeau L.J., Somerville C.C.,
 RA Spain J.C.;
 RT "Sequence Analysis and Initial Characterization of Two Isozymes of
 RT Hydroxylaminobenzene Mutase from Pseudomonas pseudocalligenes JS45.";
 RT Appl. Environ. Microbiol. 66:2965-2971(2000).

DR EMBL: AF028594; AAB94122.1; -
 SO SEQUENCE 135 AA; 14609 MW; F7502CAB7F143972 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 5; DB 2; Length 135;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 HRSRP 10
 Db 101 HRSRP 105

RESULT 131
 ID 09PGJ5 PRELIMINARY; PRT; 135 AA.
 AC 09PGJ5;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE NADH-ubiquinone oxidoreductase, NQO7 subunit.
 GN XP0305.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 OX Xylella.
 OX NCBI_TaxID=2371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9A5C;
 RX MEDLINE=20365717; PubMed=10910347;
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Britones M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrier H.,
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
 RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferio J.R.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,
 RA Garner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hohnsels J.D., Junqueira M.L., Kemper E.L., Kitaajima J.P.,
 RA Krieger J.E., Kurama E.E., Laigret F., Lambis M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madelira A.M.B.N., Madelira H.M.F., Marino C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhani A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuchino M.H.,
 RA Vallada H., Van Sluys W.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
 RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
 RL Nature 406:151-159(2000).
 DR EMBL: AE003884; AAF83116.1; -
 DR InterPro: IPR000440; Oxidored_q4.
 DR InterPro: IPR002088; PRTA.
 DR Pfam: PF00507; Oxidored_q4; 1.
 DR PROSITE: PS00904; PRTA; UNKNOWN_1.
 KM Complete proteome.
 SO SEQUENCE 135 AA; 15432 MW; DC070151F87E61BB CRC64;

Query Match
 Best Local Similarity 100.0%; Score 5; DB 16; Length 135;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPFR 7
 Db 46 LGPFR 50

OS Brucella melitensis.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=16M / ATCC 23456 / BIONTYPE 1;
RX MEDLINE=20020109; PubMed=11756688;
RA DelVecchio V.G., Kaprethel V., Redkar R.J., Patra G., Mujer C., Los T.,
RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltzman E.,
RA Savelkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-D.,
RT Haselborn R., Kyripides N., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
Brucella melitensis.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
DR EMBL: AE009625; AAL53099.1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 127 AA; 13962 MW; 6196DF62F9F467E8 CRC64;

Query Match 33.3%; Score 5; DB 16; Length 127;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 ESRAA 15
|||||
Db 46 ESRAA 50

RESULT 126
ID Q88013 PRELIMINARY; PRT; 130 AA.
AC Q88013;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE TAR protein (Transactivating regulatory protein).
GN TAR.
OS Chimpanzee immunodeficiency virus (SIV/cpz) (CIV).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11723;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=239;
RA Kestler H., Kodama T., Ringler D., Marthas M., Pedersen N.C.,
RA Lackner A., Regier D., Sehgal P., Daniel M., King N., Desrosiers R.,
RL Submitted (JUN-1990) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=239;
RX MEDLINE=90260657; PubMed=2160735;
RA Kestler H., Kodama T., Ringler D., Marthas M., Pedersen N.C.,
RA Lackner A., Regier D., Sehgal P., Daniel M., King N., Desrosiers R.,
RT "Induction of AIDS in Rhesus monkeys by molecularly cloned simian
immunodeficiency virus.";
RL Science 248:1109-1112(1990).
CC -1- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
PROMOTER (BY SIMILARITY).
DR EMBL: M33262; AAA47630.1; -;
DR InterPro: IPR001831; HIV_Tat.
DR Pfam: PF00539; Tat; 1.
DR PRINTS: PR00055; HIVTATDOMAIN.
KW Activator; Nuclear protein; RNA-binding; Transcription regulation.
SO SEQUENCE 130 AA; 14599 MW; E7870F9EE2F8BD8 CRC64;

Query Match 33.3%; Score 5; DB 15; Length 130;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 SMPES 12
|||||

Db 28 STPES 32

RESULT 127
ID Q90EX7 PRELIMINARY; PRT; 131 AA.
AC Q90EX7;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE TAR protein (Transactivating regulatory protein).
GN TAR.
OS Simian immunodeficiency virus 17E-Fr.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=160753;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97366637; PubMed=9223467;
RA Flaherty M.T., Hauer D.A., Mankowski J.E., Zink M.C., Clements J.E.,
RT "Molecular and biological characterization of a neurovirulent
molecular clone of simian immunodeficiency virus.";
RL J. Virol. 71:5790-5798(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Flaherty M.T., Hauer D.A., Mankowski J.E., Zink M.C., Clements J.E.,
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
PROMOTER (BY SIMILARITY).
DR EMBL: AY033146; AAK64190.1; -;
DR InterPro: IPR001831; HIV_Tat.
DR Pfam: PF00539; Tat; 1.
KW Activator; Nuclear protein; RNA-binding; Transcription regulation.
SO SEQUENCE 131 AA; 14800 MW; F933079380CF1D88 CRC64;

Query Match 33.3%; Score 5; DB 15; Length 131;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 SMPES 12
|||||
Db 28 STPES 32

RESULT 128
ID Q90EX2 PRELIMINARY; PRT; 131 AA.
AC Q90EX2;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE TAR protein (Transactivating regulatory protein).
GN TAR.
OS Simian immunodeficiency virus 17E-Cl.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=160754;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93331720; PubMed=8337835;
RA Anderson M.G., Hauer D., Sharma D.P., Joag S.V., Narayan O.,
RA Zink M.C., Clements J.E.,
RT "Analysis of envelope changes acquired by SIVmac239 during
neuroadaptation in rhesus macaques.";
RL Virology 195:616-626(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA Anderson M.G., Hauer D.A., Sharma D.P., Joag S.V., Narayan O.,
RA Zink M.C., Clements J.E.,
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR

Query Match 33.3%; Score 5; DB 15; Length 131;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RA Kawarabayasi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
 RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
 RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
 RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushiida N., Oguchi A.,
 RA Aoki K.-I., Masuda S., Yamaguchi M., Nishimura M., Yamagishi A.,
 RA Oshima T., Kikuchi H.;
 RT "Complete genome sequence of an aerobic thermophilic
 RT Crenarchaeon, Sulfolobus tokodaii strain7.";
 RL DNA Res. 8:123-140(2001).
 DR EMBL: AP000983; BAB5612.1; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 118 AA; 12734 MW; 5C91BACBE0104B42 CRC64;

Query Match 33.3%; Score 5; DB 17; Length 118;
 Best local Similarity 100.0%; Pred. No. 3.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STRES 12
 DB 94 STRES 98

RESULT 122
 ID 011311 PRELIMINARY; PRT; 119 AA.
 AC 011311;
 DT 01-JUL-1997 (TREMBLrel. 04, Created)
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE H2-14.2 Protein (Fragment).
 GN H2-14.2.
 OS Molluscum contagiosum virus subtype 1 (MCV1).
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Molluscipoxvirus.
 NC NCB1_TaxID=10280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Moratilla M., Agromayor M., Nunez A., Funes J.M., Varas A.J.,
 RA Lopez-Esteban J.L., Esteban M., Martin-Gallardo A.;
 RT "A Random DNA Sequencing, Computer-Based Approach for the Generation
 RT of a Gene Map of Molluscum Contagiosum Virus.";
 RL Virus Genes 0:0-0(1997).
 DR EMBL: U86896; AAB57941.1; -
 FT NON_TER 119 119
 SQ SEQUENCE 119 AA; 12728 MW; D8603AE2D44F3C96 CRC64;

Query Match 33.3%; Score 5; DB 12; Length 119;
 Best local Similarity 100.0%; Pred. No. 3.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
 DB 58 RSTPE 62

RESULT 123
 ID 08T1Y0 PRELIMINARY; PRT; 121 AA.
 AC 08T1Y0;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Predicted protein.
 GN MA4009.
 OS Methanosarcina acetivorans.
 OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
 OC Methanosarcinaceae; Methanosarcina.
 OX NCB1_TaxID=2214;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C2A / ATCC 35395 / DSM 2834;
 DE MEDLINE=21929760; PubMed=11932238;
 RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,

RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Alnoor D., Brown A.,
 RA Allen N., Naylor P., Stange-Thomann N., DeArliano K., Johnson R.,
 RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
 RA Zimmer A., Barber R.D., Cann I., Graham D.E., Graham D.A., Guss A.M.,
 RA Hedderich R., Ingram-Smith C., Kuettinger H.C., Krzycki J.A.,
 RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
 RA Springer T.A., Umayam L.A., White O., White R.H., de Maccario E.C.,
 RA Ferry J.G., Jarrell K.F., Jing H., Maccario A.J.L., Paulsen I.,
 RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
 RA Metcalf W.W., Birren B.;
 RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
 RT and physiological diversity.";
 RL Genome Res. 12:532-542(2002).
 DR EMBL: AE011113; AAM07359.1; -
 KW Complete proteome.
 SQ SEQUENCE 121 AA; 13891 MW; 80DBEAC27ECF1F2 CRC64;

Query Match 33.3%; Score 5; DB 17; Length 121;
 Best local Similarity 100.0%; Pred. No. 3.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESR 13
 DB 61 TPESR 65

RESULT 124
 ID 092RT8 PRELIMINARY; PRT; 127 AA.
 AC 092RT8;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical protein R00757.
 GN R00757 OR SMC00812.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 NC NCB1_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RA MEDLINE=21396507; PubMed=11481430;
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
 RA Boistard P., Becker A., Boutry M., Cadieu E., Dreno S., Gloux S.,
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masny D.,
 RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger O.,
 RA Renard C., Thebaud P., Vandendol M., Weidner S., Galibert F.;
 RT "Analysis of the chromosome sequence of the legume symbiont
 RT Sinorhizobium meliloti strain 1021.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
 DR EMBL: AL591784; CAC45329.1; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 127 AA; 13561 MW; D3696F6B96975FA CRC64;

Query Match 33.3%; Score 5; DB 16; Length 127;
 Best local Similarity 100.0%; Pred. No. 3.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRRA 15
 DB 74 ESRRA 78

RESULT 125
 ID 08YEG0 PRELIMINARY; PRT; 127 AA.
 AC 08YEG0;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hypothetical cytosolic protein BMEI1918.
 GN BMEI1918.

Query Match 33.3%; Score 5; DB 12; Length 112;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRA 14
 |||||
 DB 47 PESRA 51

RESULT 118

O8XV40 PRELIMINARY: PRT; 112 AA.
 AC O8XV40:
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Probable periplasmic divalent cation tolerance protein.
 GN C07A OR RSC2991 OR R501125.
 OS Ralstonia solanacearum (Pseudomonas solanacearum).
 OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
 OC Ralstonia.
 OX NCBI_TaxId-305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-GM11000;
 RX MEDLINE-21681879; PubMed-11823852;
 RA Salenoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
 Ariat M., Billaut A., Brottier P., Camus J.C., Catolico L.,
 Chaudier M., Choisme N., Claudel-Renard C., Cunne S., Demange N.,
 Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
 Siglier P., Thebaud P., Whalen M., Winkler P., Levy M.,
 Weisenbach J., Boucher C.A.;
 RA "Genome sequence of the plant pathogen Ralstonia solanacearum.";
 RL Nature 415:497-502(2002).
 DR EMBL: AL646073; CAD16700.1; -;
 DR InterPro: IPR004323; CUFAL.
 DR Pfam: PF03091; CUFAL.1.
 KW Complete proteome.
 SQ SEQUENCE 112 AA; 11943 MW; D53F6C21F231BF96 CRC64;

Query Match 33.3%; Score 5; DB 16; Length 112;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRAA 15
 |||||
 DB 29 ESRAA 33

RESULT 119

P81170 PRELIMINARY: PRT; 113 AA.
 AC P81170:
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Putative hippurate hydrolase (EC 3.5.1.32) (BENZOXGLYCINE
 DE amidohydrolase) (Fragment).
 OS Trifolium repens (Creeping white clover).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Trifolium.
 OX NCBI_TaxId-3899;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Koberstaedt A., Lenz M., Reley J.;
 RL Submitted (NOV-1992) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP IDENTIFICATION, AND CONCEPTUAL TRANSLATION.
 RA Rawlings N.D.;
 RL Unpublished observations (JAN-1998).
 CC -I- CATALYTIC ACTIVITY: HIPPURATE + H(2)O = BENZOATE + GLYCINE.

CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M40, ALSO KNOWN AS THE
 CC AMA/HIPO/HYUC FAMILY OF HYDROLASES.
 CC -I- CAUTION: THIS IS A CONCEPTUAL TRANSLATION.
 DR EMBL: X68950; -, NOT_ANNOTATED_CDS.

KW Hypothetical protein; Hydrolase; Zinc.
 FT METAL 63 63 ZINC (POTENTIAL).
 FT METAL 65 65 ZINC (POTENTIAL).
 FT NON_TER 113 113
 SQ SEQUENCE 113 AA; 11341 MW; 6AE43DC29C96C3DD CRC64;

Query Match 33.3%; Score 5; DB 10; Length 113;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 PRST 9
 |||||
 DB 76 PRST 80

RESULT 120

O923U4 PRELIMINARY: PRT; 117 AA.
 AC O923U4:
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE TRP4 (Fragment).
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystriognathii; Cavidae; Cavia.
 OX NCBI_TaxId-10141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ong H.L., Breerton H.M., Harland M.L., Barritt G.J.;
 RT "Expression of mRNA encoding trp cation channels in guinea pig airway
 RT smooth muscle cells.";
 RL Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF395786; AAK84033.1; -;
 DR InterPro: IPR002111; Cat_channel_TrpL.
 DR InterPro: IPR000636; M+channel_nlg.
 DR InterPro: IPR001865; Ribosomal_S2.
 DR Pfam: PF00520; Ion_trans.1.
 DR PROSITE: PS00962; RIBOSOMAL_S2_1; UNKNOWN_1.
 FT NON_TER 1 1
 FT NON_TER 117 117
 SQ SEQUENCE 117 AA; 13211 MW; 9ABC93700F2E606C CRC64;

Query Match 33.3%; Score 5; DB 11; Length 117;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGP 5
 |||||
 DB 28 SHLGP 32

RESULT 121

O974P5 PRELIMINARY: PRT; 118 AA.
 AC O974P5:
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Hypothetical protein ST0614.
 GN ST0614.
 OS Sulfolobus tokodaii.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 OX NCBI_TaxId-111955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-JCM 10545 / 7;
 RX PubMed-11572479;

DB 32 PESRA 36

RESULT 114

08VYN2 PRELIMINARY; PRT; 110 AA.
AC 08VYN2;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Putative transposase.
OS Vibrrio Cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MO10; TRANSPONSON-SXT ELEMENT;
RX MEDLINE=21485931; PubMed=11600347;
RA Hochhut B., Lotfi Y., Mazel D., Faruque S.M., Woodgate R., Waldor M.K.;
RT "Molecular Analysis of Antibiotic Resistance Gene Clusters in *Vibrio cholerae* O139 and O1 SXT Constructions."
RL Antimicrob. Agents Chemother. 45:2991-3000(2001).
DR EMBL; AY034138; AAK64585.1; -
SQ SEQUENCE 110 AA; 12071 MW; 4AB78ED592E93216 CRC64;

Query Match 33.3%; Score 5; DB 2; Length 110;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRA 14
DB 95 PESRA 99

RESULT 115

09EPT7 PRELIMINARY; PRT; 110 AA.
AC 09EPT7;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Ribosomal protein S10.
GN RPS10.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20548717; PubMed=11099041;
RA Adams K.L., Daley D.O., Qiu Y.L., Whelan J., Palmer J.D.;
RT "Repeated, recent and diverse transfers of a mitochondrial gene to the RT nucleus in flowering plants."
RL Nature 408:354-357(2000).
DR EMBL; AF302137; AAG37233.1; -
DR InterPro; IPR001848; Ribosomal_S10.
DR Pfam; PF00338; Ribosomal_S10; 1.
DR PRINTS; PR00971; RIBOSOMAL_S10.
DR ProDom; PD001272; Ribosomal_S10; 1.
SQ SEQUENCE 110 AA; 12936 MW; F75BFDBA6DCFA51 CRC64;

Query Match 33.3%; Score 5; DB 10; Length 110;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRA 14
DB 32 PESRA 36

RESULT 116

09D503 PRELIMINARY; PRT; 112 AA.
AC 09D503;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE 493053B18RIK protein.
GN 493053B18RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TESTIS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawal J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K., Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schriml L.M., Staib F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmink L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S., Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL; AK015952; BAB30049.1; -
DR MGD; MGI:1922418; 493053B18RIK.
SQ SEQUENCE 112 AA; 12360 MW; D31828B0D55BEC61 CRC64;

Query Match 33.3%; Score 5; DB 11; Length 112;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 TPESR 13
DB 28 TPESR 32

RESULT 117

09YR60 PRELIMINARY; PRT; 112 AA.
AC 09YR60;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Nonstructural protein 1 (Fragment).
GN NS1.
OS Aleutian mink disease virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=28314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A2;
RA Olofsson A., Mittelholzer C., Treiberg Berndtsson L., Lind L., Mejerland T., Belak S.;
RT "Unusual, high genetic diversity of Aleutian mink disease virus."
RL J. Clin. Microbiol. 39:4145-4149(1999).
DR EMBL; AF107627; AAD04593.1; -
FT NON_TER 1 112
SQ SEQUENCE 112 AA; 13103 MW; A6950872B9ED1790 CRC64;

DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Peptide Y.
GN PY.
OS Parallichthys olivaceus (Flounder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorphi; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectoidae; Parallichthyidae; Parallichthys.
OX NCBI_TaxId=8255;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Kurokawa T., Suzuki T.;
RT "Development of neuropeptide Y related peptides in the digestive
RT organs during the larval stage of Japanese flounder, Parallichthys
RT olivaceus.";
RL Submitted (Feb-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB055213; BAB62411.1; -
DR InterPro: IPR001955; Pancreatic_hormn.
DR Pfam: PF00159; hormone3.1.
DR ProDom: PD001267; Pancreatic_hormn.1.
DR PROSITE: PS00265; PANCREATIC_HORMONE_2; 1. UNKNOWN_1.
DR PROSITE: PS50276; PANCREATIC_HORMONE_2; 1.
SQ SEQUENCE 97 AA; 10929 MW; 8703DD3933FE175 CRC64;

Query Match 33.3%; Score 5; DB 13; Length 97;
Best Local Similarity 100.0%; Pred. No. 2,8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
DB 67 RSTPE 71

RESULT 111
O82DG2 PRELIMINARY; PRT; 105 AA.
AC O82DG2;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical protein YP02606.
GN YP02606.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
OX NCBI_TaxId=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CO-92 / BIOVAR ORIENTALIS;
RX MEDLINE-21470413; PubMed-11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebatina M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Fellwell T., Hamlin N., Holtroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA Slimmons M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
DR EMBL; AJ414153; CAC92849.1; -
DR InterPro: IPR003456; DUF143.
DR InterPro: IPR004394; Iojap.
DR Pfam: PF02410; DUF143; 1.
DR TIGRFAMs: TIGR00090; Iojap; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 105 AA; 11605 MW; A8A87A37D2ACBCE CRC64;

Query Match 33.3%; Score 5; DB 16; Length 105;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRAA 15
DB 58 ESRAA 62

RESULT 112
O99494 PRELIMINARY; PRT; 106 AA.
AC O99494;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE DAT1 protein (Fragment).
GN DAT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BLOOD;
RA Lerman C., Caporaso N.E., Main D., Audrain J., Bowman E.D.,
RA Lockshin B., Boyd N.R., Shields P.G.;
RT "Association of dopamine transporter (DAT1) and DRD2 receptor genes
RT with smoking.";
RL Am. J. Hum. Genet. 0:0-0(0).
DR EMBL; Y10141; CAA71229.1; -
FT NON_TER 1 1
FT NON_TER 106 106
SQ SEQUENCE 106 AA; 10953 MW; 2AD345DF1966B20C CRC64;

Query Match 33.3%; Score 5; DB 4; Length 106;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 GPHRS 8
DB 31 GPHRS 35

RESULT 113
O9FUT8 PRELIMINARY; PRT; 109 AA.
AC O9FUT8;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Ribosomal protein S10.
DE Zea mays (Maize).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogonaceae; Zea.
OX NCBI_TaxId=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-20548717; PubMed-11099041;
RA Adams K.L., Daley D.O., Olin Y.L., Whelan J., Palmer J.D.;
RT "Repeated, recent and diverse transfers of a mitochondrial gene to the
RT nucleus in flowering plants.";
RL Nature 408:354-357(2000).
DR EMBL; AF287341; AAG32319.1; -
DR InterPro: IPR001848; Ribosomal_S10.
DR Pfam: PF00338; Ribosomal_S10; 1.
DR PRINTS: PR00971; RIBOSOMALS10.
DR ProDom: PD001272; Ribosomal_S10; 1.
SQ SEQUENCE 109 AA; 12684 MW; 7EB23733E4CFBFC6 CRC64;

Query Match 33.3%; Score 5; DB 10; Length 109;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRA 14

RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
 DR EMBL: AE009584; AAL52655.1; -
 KW Hypothetical protein; Complete proteome.
 SO SEQUENCE 87 AA; 9072 MW; 4D19305B02B41281 CRC64;

Query Match 33.3%; Score 5; DB 16; Length 87;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRA 14
 DB 77 PESRA 81

RESULT 106

O9N001 PRELIMINARY; PRT; 91 AA.
 AC O9N001;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DE Hypothetical 9.6 kDa protein.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Macaca.
 OX NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=CEREBELLUM CORTEX;
 RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
 RA Suzuki Y., Sugano S., Hashimoto K.;
 RT "Isolation of full-length cDNA clones from macaque brain cDNA
 RT libraries."
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB046646; BAB03564.1; -
 KW Hypothetical protein.
 SO SEQUENCE 91 AA; 9574 MW; BCAA305E8E4E53FB CRC64;

Query Match 33.3%; Score 5; DB 6; Length 91;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGP 5
 DB 48 SHLGP 52

RESULT 107

O50101 PRELIMINARY; PRT; 95 AA.
 AC O50101;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DE Hypothetical 10.2 kDa protein U650J.
 GN U650J.
 OS Mycobacterium leprae.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1769;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Smith D.R., Robison K.;
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U15184; AAA63068.1; -
 KW Hypothetical protein.
 SO SEQUENCE 95 AA; 10229 MW; 1037F851EFA214C2 CRC64;

Query Match 33.3%; Score 5; DB 2; Length 95;
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGP 5
 DB 84 SHLGP 88

RESULT 108

O9N8N8 PRELIMINARY; PRT; 96 AA.
 AC O9N8N8;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DE Possible a580r protein.
 GN CHRI.237.
 OS Trypanosoma brucei.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 OX NCBI_TaxID=5691;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TREU927;
 RA Hall N., Bowman S., Quail M., Ivens A.C., Kay M.P., Bray-Allen S.,
 RA Leonard N.J., Clark L.N., Harris B.R., Melville S., Lawson D.,
 RA Gerrard C., Rajandream M.A., Barrell B.G.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL359782; CAB95505.1; -
 SO SEQUENCE 96 AA; 11222 MW; 3CC038ED637763B1 CRC64;

Query Match 33.3%; Score 5; DB 5; Length 96;
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRAA 15
 DB 37 ESRAA 41

RESULT 109

O8SC82 PRELIMINARY; PRT; 96 AA.
 AC O8SC82;
 DT 01-JUN-2002 (TReMBLrel. 21, Created)
 DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
 DE Hypothetical 10.7 kDa protein.
 OS Stx2 converting bacteriophage I.
 OC Viruses.
 OX NCBI_TaxID=180816;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-STX2 PHAGE-I;
 RA Sato T., Shimizu T., Matarai M., Kobayashi M., Kano S., Hamabata T.,
 RA Yamasaki S., Takeda Y.;
 RT "Genomic sequence of Shiga toxin 2-converting phage isolated from
 RT Escherichia coli O157:H7 Okayama strain and comparison with other
 RT Shiga toxin 2-converting phages."
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP004402; BAB87902.1; -
 KW Hypothetical protein.
 SO SEQUENCE 96 AA; 10747 MW; 567C7F96EA9A70F6 CRC64;

Query Match 33.3%; Score 5; DB 9; Length 96;
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PHRST 9
 DB 77 PHRST 81

RESULT 110

O90WF2 PRELIMINARY; PRT; 97 AA.
 AC O90WF2;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DE Hypothetical 9.6 kDa protein.
 GN U650J.
 OS Mycobacterium leprae.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1769;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Smith D.R., Robison K.;
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U15184; AAA63068.1; -
 KW Hypothetical protein.
 SO SEQUENCE 95 AA; 10229 MW; 1037F851EFA214C2 CRC64;

DR Prodom: PD000328; MHC_II_beta.1.
 KW Glycoprotein; MHC II; Transmembrane.
 FT NON_TER 1
 FT NON_TER 83
 SQ SEQUENCE 83 AA; 9733 MW; 57AE6E9E0EBB840 CRC64;

Query Match 33.3%; Score 5; DB 7; Length 83;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0;

OY 11 ESRAA 15
 DB 62 ESRAA 66

RESULT 102
 019227 PRELIMINARY; PRT; 83 AA.

AC 019227;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE MHC II DR-beta 1 chain (Fragment).
 GN DRB1.
 OS Rangifer tarandus (Reindeer) (Caribou).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidae;
 CC Cervidae; Odocoileinae; Rangifer.
 NC NCB1_TaxID=9870;

RN [1]
 RP SEQUENCE FROM N.A.
 RA Mikko S., Roed K., Schmutz S., Andersson L.;
 RT "Genetic diversity at major histocompatibility complex DRB loci in
 some domesticated and wild ruminant species";
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF012723; AAB6616.1; -.
 DR HSSP: P13758; 1DLH.
 DR InterPro: IPR000353; MHC_II_beta.
 DR Pfam: PF00969; MHC_II_beta.1.
 DR Prodom: PD000328; MHC_II_beta.1.
 KW Glycoprotein; MHC II; Transmembrane.

FT NON_TER 1
 FT NON_TER 83
 SQ SEQUENCE 83 AA; 9612 MW; 4FA02D4C788A896 CRC64;

Query Match 33.3%; Score 5; DB 7; Length 83;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0;

OY 11 ESRAA 15
 DB 62 ESRAA 66

RESULT 103
 095MA6 PRELIMINARY; PRT; 85 AA.

AC 095MA6;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Glucokinase (Fragment).
 OS Equus caballus (Horse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 NC NCB1_TaxID=9796;

RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-2131992; PubMed-11421942;
 RA Shubtowski D.M., Venta P.J., Douglass C.L., Zhou R.-X., Ewart S.L.;
 RT "Polymorphism identification within 50 equine gene-specific sequence
 tagged sites";
 RL Anim. Genet. 32:78-78(2001).

DR EMBL: AY008787; AAC43039.1; -.
 DR InterPro: IPR001312; Hexokinase.
 DR Pfam: PF00349; hexokinase.1.
 DR Prodom: PD001109; Hexokinase.1.
 KW Kinase.
 FT NON_TER 1
 FT NON_TER 85
 SQ SEQUENCE 85 AA; 9640 MW; 01000FA744606590 CRC64;

Query Match 33.3%; Score 5; DB 6; Length 85;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0;

OY 7 RSTPE 11
 DB 40 RSTPE 44

RESULT 104
 096A26 PRELIMINARY; PRT; 87 AA.

AC 096A26;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Similar to interferon stimulated gene (20kp).
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NC NCB1_TaxID=9606;

RN [1]
 RP SEQUENCE FROM N.A.
 RA Tissue-BRAIN;
 RC Strausberg R.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC016341; AAH16341.1; -.
 DR InterPro: IPR000520; Exonuclease.
 DR Pfam: PF00929; Exonuclease.1.

DR Pfam: PF00929; Exonuclease.1.
 SQ SEQUENCE 87 AA; 9298 MW; EAF8FF51046D16A6 CRC64;

Query Match 33.3%; Score 5; DB 4; Length 87;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0;

OY 3 LGPHR 7
 DB 17 LGPHR 21

RESULT 105
 08YFP6 PRELIMINARY; PRT; 87 AA.

AC 08YFP6;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hypothetical protein BMEI1474.
 GN BMEI1474.
 OS Brucella melitensis.
 CC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 CC Brucellaceae; Brucella.
 NC NCB1_TaxID=29459;

RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=16M / ATCC 23456 / B10TYPE 1;
 RA MEDLINE-20020109; PubMed-11756688;
 RA DelVecchio V.G., Kapral V., Redkar R.J., Patra G., Mufer C., Los T.,
 RA Ivanova N., Anderson I., Bhattacharya A., Lykdis A., Reznik G.,
 RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,
 RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
 RA Haselkorn R., Kypides N., Overbeek R.;
 RT "The genome sequence of the facultative intracellular pathogen
 Brucella melitensis";

```

SQ SEQUENCE 78 AA; 9114 MW; 0DEBD412F261B3782 CRC64;
Query Match
Best Local Similarity 100.0%; Score 5; DB 7; Length 78;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRAA 15
    |||||
Db 62 ESRAA 66

RESULT 98
Q8WM08
ID 08WM08 PRELIMINARY; PRT; 78 AA.
AC 08WM08;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DE MHC class II DR alpha 1 domain (Fragment).
GN RAPA-DRB1.
OS Rangifer tarandus (Reindeer) (Caribou).
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidae;
OC Cervidae; Odocoileinae; Rangifer.
OX NCBI_TaxID=9870;
RN [1]
RP SEQUENCE FROM N.A.
RA Wei Z., Hupp G.;
RT "MHC variation in reindeer and caribou.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF458948; AAL67477.1; -.
DR InterPro; IPR000353; MHC_II_beta.
DR Pfam; PF00969; MHC_II_beta; 1.
DR ProDom; PD000328; MHC_II_beta; 1.
FT NON_TER 1
FT NON_TER 78
SQ SEQUENCE 78 AA; 9307 MW; 5C6B403E66508FB0 CRC64;

Query Match
Best Local Similarity 100.0%; Score 5; DB 7; Length 78;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRAA 15
    |||||
Db 62 ESRAA 66

RESULT 99
019220 PRELIMINARY; PRT; 83 AA.
ID 019220;
AC 019220;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MHC II DR-beta 1 chain (Fragment).
GN DRB1.
OS Rangifer tarandus (Reindeer) (Caribou).
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidae;
OC Cervidae; Odocoileinae; Rangifer.
OX NCBI_TaxID=9870;
RN [1]
RP SEQUENCE FROM N.A.
RA Mikko S., Roed K., Schmutz S., Andersson L.;
RT "Genetic diversity at major histocompatibility complex DRB loci in
some domesticated and wild ruminant species.";
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF012716; AAB66609.1; -.
DR InterPro; IPR000353; MHC_II_beta.
DR Pfam; PF00969; MHC_II_beta; 1.
DR ProDom; PD000328; MHC_II_beta; 1.
DR Glycoprotein; MHC II; Transmembrane.
FT NON_TER 1

```

```

FT NON_TER 83
SQ SEQUENCE 83 AA; 9745 MW; 07B171EAD0D8AF4B CRC64;
Query Match
Best Local Similarity 100.0%; Score 5; DB 7; Length 83;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRAA 15
    |||||
Db 62 ESRAA 66

RESULT 100
019221 PRELIMINARY; PRT; 83 AA.
ID 019221;
AC 019221;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MHC II DR-beta 1 chain (Fragment).
GN DRB1.
OS Rangifer tarandus (Reindeer) (Caribou).
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidae;
OC Cervidae; Odocoileinae; Rangifer.
OX NCBI_TaxID=9870;
RN [1]
RP SEQUENCE FROM N.A.
RA Mikko S., Roed K., Schmutz S., Andersson L.;
RT "Genetic diversity at major histocompatibility complex DRB loci in
some domesticated and wild ruminant species.";
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF012717; AAB66610.1; -.
DR InterPro; IPR000353; MHC_II_beta.
DR Pfam; PF00969; MHC_II_beta; 1.
DR ProDom; PD000328; MHC_II_beta; 1.
DR Glycoprotein; MHC II; Transmembrane.
FT NON_TER 1
FT NON_TER 83
SQ SEQUENCE 83 AA; 9815 MW; 57BF9FEAD167E840 CRC64;

Query Match
Best Local Similarity 100.0%; Score 5; DB 7; Length 83;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRAA 15
    |||||
Db 62 ESRAA 66

RESULT 101
019222 PRELIMINARY; PRT; 83 AA.
ID 019222;
AC 019222;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MHC II DR-beta 1 chain (Fragment).
GN DRB1.
OS Rangifer tarandus (Reindeer) (Caribou).
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidae;
OC Cervidae; Odocoileinae; Rangifer.
OX NCBI_TaxID=9870;
RN [1]
RP SEQUENCE FROM N.A.
RA Mikko S., Roed K., Schmutz S., Andersson L.;
RT "Genetic diversity at major histocompatibility complex DRB loci in
some domesticated and wild ruminant species.";
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF012718; AAB66611.1; -.
DR InterPro; IPR000353; MHC_II_beta.
DR Pfam; PF00969; MHC_II_beta; 1.

```

Query Match 33.3%; Score 5; DB 16; Length 69;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRAA 15
Db 22 ESRAA 26

RESULT 94

O8XVKO PRELIMINARY; PRT; 71 AA.
AC O8XVKO;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Hypothetical protein RSC2830.
GN RSC2830 OR RSC0275.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM11000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Christine N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie W., Moisan A., Robert C., Saurin W., Schlex T.,
RA Sigulier P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL; AL646072; CADI6537.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 71 AA; 7726 MW; BB8116a6554BB043 CRC64;

Query Match 33.3%; Score 5; DB 16; Length 71;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESR 13
Db 18 TPESR 22

RESULT 95

O9E796 PRELIMINARY; PRT; 72 AA.
AC O9E796;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE Core protein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D2;
RX MEDLINE=20472685; PubMed=11015367;
RA Ross R.S., Viazov S.O., Holtzer C.D., Beyou A., Monnet A., Mazure C.,
RA Roggendorf M.;
RT "Genotyping of hepatitis C virus isolates using CLIP sequencing.";
RL J. Clin. Microbiol. 38:3581-3584(2000).
DR EMBL; AF233719; AAG24348.1; -
FT NON_TER 1
FT NON_TER 72
FT NON_TER 72
SQ SEQUENCE 72 AA; 7942 MW; B0D30C642AB6818C CRC64;

Query Match 33.3%; Score 5; DB 12; Length 72;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHR 7
Db 67 LGPHR 71

RESULT 96

O9YE07 PRELIMINARY; PRT; 76 AA.
AC O9YE07;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Hypothetical protein APE5034.
GN APE5034.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcaceae;
OC Desulfurococcaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankaei A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000060; BAA9740.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 76 AA; 8245 MW; 9C8C49D7D4E8F559 CRC64;

Query Match 33.3%; Score 5; DB 17; Length 76;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
Db 11 RSTPE 15

RESULT 97

O8WM11 PRELIMINARY; PRT; 78 AA.
AC O8WM11;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE MHC class II DR alpha 1 domain (Fragment).
GN RAPA-DB1.
OS Rangifer tarandus (Reindeer) (Caribou).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidae;
OC Cervidae; Odocoileinae; Rangifer.
OX NCBI_TaxID=9870;
RN [1]
RP SEQUENCE FROM N.A.
RC Wei Z., Happ G.;
RT "MHC variation in reindeer and caribou.";
RT Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF458945; AAL67474.1; -
DR InterPro; IPR000353; MHC_II_beta.
DR Pfam; PF00969; MHC_II_beta.1.
DR ProDom; PD000328; MHC_II_beta.1.
FT NON_TER 1
FT NON_TER 78
FT NON_TER 78
SQ SEQUENCE 78 AA; 8245 MW; 9C8C49D7D4E8F559 CRC64;

RL Virology 167:485-496(1988).
 RN [15]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-92196996; PubMed-1549908;
 RA Handermann M., Schmitzler P., Rosen-Wolff A., Raab K., Sonntag K.C.,
 RT Dairai G.;
 RL "Identification and mapping of origins of DNA replication within the
 RT DNA sequences of the genome of insect iridescent virus type 6.";
 RL Virus Genes 6:19-32(1992).
 RN [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-93118242; PubMed-1475907;
 RA Sonntag K.C., Dairai G.;
 RT "Characterization of the third origin of DNA replication of the genome
 RT of insect iridescent virus type 6.";
 RL Virus Genes 6:333-342(1992).
 RN [7]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-93260401; PubMed-8492091;
 RA Stowasser R., Raab K., Schmitzler P., Janssen W., Dairai G.;
 RT "Identification of the gene encoding the major capsid protein of
 RT insect iridescent virus type 6 by polymerase chain reaction.";
 RL J. Gen. Virol. 74:873-879(1993).
 RN [8]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-94167241; PubMed-8121799;
 RA Schmitzler P., Hug M., Handermann M., Janssen W., Koonin E.V.,
 RA Dellus H., Dairai G.;
 RT "Identification of genes encoding zinc finger proteins, non-histone
 RT chromosomal HMG protein homologue, and a putative GTP phosphohydrolase
 RT in the genome of Chilo iridescent virus.";
 RL Nucleic Acids Res. 22:158-166(1994).
 RN [9]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-94353641; PubMed-8073636;
 RA Sonntag K.C., Schmitzler P., Koonin E.V., Dairai G.;
 RT "Chilo iridescent virus encodes a putative helicase belonging to a
 RT distinct family within the 'DEAD/H' superfamily: implications for the
 RT evolution of large DNA viruses.";
 RL Virus Genes 8:151-158(1994).
 RN [10]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-95213160; PubMed-7698884;
 RA Sonntag K.C., Schmitzler P., Janssen W., Dairai G.;
 RT "Identification of the primary structure and the coding capacity of
 RT the genome of insect iridescent virus type 6 between the genome
 RT coordinates 0.310 and 0.347 (7990 bp).";
 RL Intervirology 37:287-297(1994).
 RN [11]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-94292906; PubMed-8021587;
 RA Schmitzler P., Sonntag K.C., Muller M., Janssen W., Bugett J.J.,
 RA Koonin E.V., Dairai G.;
 RT "Insect iridescent virus type 6 encodes a polypeptide related to the
 RT largest subunit of eukaryotic RNA polymerase II.";
 RL J. Gen. Virol. 75:1557-1567(1994).
 RN [12]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-96141693; PubMed-9482589;
 RA Bahr U., Tidona C.A., Dairai G.;
 RT "The DNA sequence of Chilo iridescent virus between the genome
 RT coordinates 0.101 and 0.391; similarities in coding strategy between
 RT insect and vertebrate Iridoviruses.";
 RL Virus Genes 15:235-245(1997).
 RN [13]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-99125223; PubMed-9926400;
 RA Muller K., Tidona C.A., Bahr U., Dairai G.;
 RT "Identification of a thymidylate synthase gene within the genome of
 RT Chilo iridescent virus.";
 RL Virus Genes 17:243-258(1998).
 RN [14]
 RP SEQUENCE FROM N.A.

RX MEDLINE-99383793; PubMed-10456793;
 RA Muller K., Tidona C.A., Dairai G.;
 RT "Identification of a gene cluster within the genome of Chilo
 RT iridescent virus encoding enzymes involved in viral DNA replication
 RT and processing.";
 RL Virus Genes 18:243-264(1999).
 RN [15]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-21342589; PubMed-11448171;
 RA Jakob N.J., Muller K., Bahr U., Dairai G.;
 RT "Analysis of the first Complete DNA Sequence of an Invertebrate
 RT Iridovirus: Coding Strategy of the Genome of Chilo Iridescent Virus.";
 RL Virology 286:182-196(2001).
 RN [16]
 RP SEQUENCE FROM N.A.
 RA Jakob N.J., Mueller K., Bahr U., Dairai G.;
 RL Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF03741; AAK8214.1; -
 SQ SEQUENCE 62 AA; 6641 MW; 27A92D06DA98F0E9 CRC64;
 Query Match 33.3%; Score 5; DB 12; Length 62;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 9 TPESR 13
 DB 31 TPESR 35
 RESULT 93
 ID O8XBP2 PRELIMINARY; PRT; 69 AA.
 AC O8XBP2;
 DT 01-MAR-2002 (TRENBLREL. 20, Created)
 DT 01-MAR-2002 (TRENBLREL. 20, Last sequence update)
 DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
 DE yber protein.
 GN YBER OR Z0783 OR EGS0675.
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OC NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-O157:H7 / EDI933 / ATCC 700927;
 RX MEDLINE-21074935; PubMed-11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
 RL Nature 409:529-533(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-O157:H7 / RIMD 0509952;
 RX MEDLINE-21156231; PubMed-11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 RA Kihara S., Shiba T., Hattori M., Shinozawa H.;
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22(2001).
 DR EMBL: AE005242; AAC54971.1; -
 DR EMBL: AP002552; BAB34098.1; -
 DR InterPro: IPR003456; DUF143.
 DR InterPro: IPR004394; Iofap.
 DR Pfam: PF02410; DUF143; 1.
 DR TIGRFAWS: TIGR00090; Iofap; 1.
 KW Complete proteome.
 SQ SEQUENCE 69 AA; 7678 MW; A470FC7D9207E12 CRC64;

```

Q939W5          PRELIMINARY;      PRT;      55 AA.
ID Q939W5
AC Q939W5
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Transposase protein (Fragment).
GN TRANSPOSASE.
OS Aeromonas salmonicida.
OC Bacteria; Proteobacteria; gamma subdivision; Aeromonadaceae;
OC Aeromonas.
OX NCBI_TaxID=645;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=92/2298;
RA Sorum H., Kiekgg O., Alveheim K.;
RT "A novel tetracycline resistance determinant, Tet 31, on a
RL transferable R-plasmid from the fish pathogen Aeromonas salmonicida."
DR EMBL; AJ250203; CAC80725.1; -.
KW Plasmid.
FT NON_TER
SQ SEQUENCE 55 AA; 5815 MW; E33F306121E0B0F4 CRC64;

Query Match          33.3%; Score 5; DB 2; Length 55;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 PESRA 14
Db 40 PESRA 44

RESULT 90
ID Q9WZ80          PRELIMINARY;      PRT;      56 AA.
AC Q9WZ80;
DT 01-NOV-1999 (TREMblrel. 12, Created)
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Hypothetical protein TM0611.
GN TM0611.
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RA MEDLINE=9287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linner K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima.";
RL Nature 393:323-329(1998).
DR EMBL; AE001735; AAD35696.1; -.
RA TIGR; TM0611; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 56 AA; 6468 MW; 346F85EB61A078A0 CRC64;

Query Match          33.3%; Score 5; DB 16; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 TPESR 13
Db 22 TPESR 26

```

```

RESULT 91
ID Q14274          PRELIMINARY;      PRT;      62 AA.
AC Q14274
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE ERV8 protein (Fragment).
GN ERV8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA MEDLINE=87036922; PubMed=3021993;
RA Ono M., Yasunaga T., Miyata T., Ushikubo H.;
RT "Nucleotide sequence of human endogenous retrovirus genome related to
RT the mouse mammary tumor virus genome."
RL J. Virol. 60:589-598(1986).
DR EMBL; K03499; AAA8034.1; -.
FT NON_TER
SQ SEQUENCE 62 AA; 6821 MW; 4729AC54BCE5110 CRC64;

Query Match          33.3%; Score 5; DB 4; Length 62;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 ESRAA 15
Db 26 ESRAA 30

RESULT 92
ID Q91FP4          PRELIMINARY;      PRT;      62 AA.
AC Q91FP4;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE 280R.
OS Chilo iridescent virus (CIV) (Insect iridescent virus type 6).
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.
OX NCBI_TaxID=10488;
RN [1]
RP SEQUENCE FROM N.A.
RA Delius H., Darai G., Fluegel R.M.;
RT "DNA analysis of insect iridescent virus 6: evidence for circular
RT permutation and terminal redundancy.";
RL J. Virol. 49:609-614(1984).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=86174607; PubMed=3959991;
RA Losbacher de Ruiz H., Gelderblom H., Hofmann W., Darai G.;
RT "Insect iridescent virus type 6 induced toxic degenerative hepatitis
RT in mice.";
RL Med. Microbiol. Immunol. 175:43-53(1986).
RN [3]
RP SEQUENCE FROM N.A.
RA MEDLINE=87321126; PubMed=2820141;
RA Schmitzler P., Soltau J.B., Fischer M., Reitsner H., Scholz J.,
RA Delius H., Darai G.;
RT "Molecular cloning and physical mapping of the genome of insect
RT iridescent virus type 6: further evidence for circular permutation of
RT the viral genome.";
RL Virology 160:66-74(1987).
RN [4]
RP SEQUENCE FROM N.A.
RA MEDLINE=89073752; PubMed=3201750;
RA Fischer M., Schmitzler P., Delius H., Darai G.;
RT "Identification and characterization of the repetitive DNA element in
RT the genome of insect iridescent virus type 6.";

```

```
DE | Tax protein (Fragment).
GN |
NC |
OS | Human T-cell leukemia virus type II (HTLV-II).
OC | Viruses; Retrovird viruses; Retroviridae; Deltaretrovirus.
OX | NCBITaxId=11909;
RN | [1]
RP | SEQUENCE FROM N.A.
RC | STRAIN=YL4;
RA | Leon-Porte M., Echeverria de Perez G., Blanco N., Hengst J., Dube S.,
RI | Love J., Polasz B.;
RT | "Endemic infection with human T-cell leukemia/lymphoma virus type IIB
RL | in Venezuelan Indians: molecular characterization.";
DR | EMBL; AF005399; AAB87757.1; -
FT | NON_TER 1 1
SQ | SEQUENCE 39 AA; 4417 MW; 26A991DD18DD65CC CRC64;

Query Match
Best Local Similarity 33.3%; Score 5; DB 15; Length 39;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HLGP 6
DB 35 HLGP 39

RESULT 86
ID 082210 PRELIMINARY; PRT; 39 AA.
AC 082210;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Tax protein (Fragment).
GN TAX.
OS Human T-lymphotropic virus 1.
OC Viruses; Retrovird viruses; Retroviridae; Deltaretrovirus.
OX NCBITaxId=11908;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSHR-1;
RA Yanagihara R., Saitou N., Nerurkar V.R., Song K.J., Bastian I.,
RI Frenchini G., Gajdusek D.C.;
RT "Molecular phylogeny and dissemination of human T-cell lymphotropic
RT virus type I viewed within the context of primate evolution and human
RT migration.";
RL Cell. Mol. Biol. 0:0-0(1995).
DR EMBL; U12121; AAA99642.1; -
FT NON_TER 1 1
SQ | SEQUENCE 39 AA; 4425 MW; 23A6CB4626765F9B CRC64;

Query Match
Best Local Similarity 33.3%; Score 5; DB 15; Length 39;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HLGP 6
DB 35 HLGP 39

RESULT 87
ID 082239 PRELIMINARY; PRT; 39 AA.
AC 082239;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE Tax protein (Fragment).
GN TAX.
OS Human T-cell leukemia virus type II (HTLV-II).
OC Viruses; Retrovird viruses; Retroviridae; Deltaretrovirus.
```

```
OX NCBITaxId=11909;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SUBTYPE B;
RX MEDLINE=96078239; PubMed=7576938;
RA Bolton W.V., Kenrick K.G., Dwyer D.E., Cunningham A.L., Wylie B.R.,
RI Saksena N.K.;
RT "Partial nucleotide sequence analysis of the first case of human T
RT lymphotropic virus type II from Australia.";
RL AIDS Res. Hum. Retroviruses 11:765-767(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SUBTYPE B;
RA Bryant B.W.;
RL Submitted (May-1995) to the EMBL/Genbank/DBJ databases.
DR EMBL; U25957; AAA91015.1; -
FT NON_TER 1 1
SQ | SEQUENCE 39 AA; 4417 MW; 26A991DD18DD65CC CRC64;

Query Match
Best Local Similarity 33.3%; Score 5; DB 15; Length 39;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HLGP 6
DB 35 HLGP 39

RESULT 88
ID 0929G4 PRELIMINARY; PRT; 51 AA.
AC 0929G4;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical protein Cpn0006.
GN Cpn0006 OR Cpn0006.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBITaxId=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kalnan S., Mitchell W., Marathe R., Lammell C., Fan J., Hyman R.W.,
RI Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RI Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
DR EMBL; AE001585; AAD18164.1; -
DR EMBL; AP002545; BAA98216.1; -
KW Hypothetical protein; Complete proteome.
SQ | SEQUENCE 51 AA; 5804 MW; 8229BE6536FCFAE CRC64;

Query Match
Best Local Similarity 33.3%; Score 5; DB 16; Length 51;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPE 11
DB 29 RSTPE 33

RESULT 89
```

014559
ID 014559 PRELIMINARY; PRT; 39 AA.
AC 014559;
DT 01-JAN-1998 (TRENBLREL. 05, Created)
DT 01-JAN-1998 (TRENBLREL. 05, Last sequence update)
DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)
DE HPX-2 protein (Fragment).
GN HPX-2.
OS Homo sapiens (Human).
OC Eumariota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BONE MARROW;
RX MEDLINE-94314219; PubMed-7518789;
RA Moretti P.A.B., Simmons P., Thomas P., Haylock D., Rathen P.,
RA Vadas M., D'Andrea R.;
RT "Identification of homeobox genes expressed in human haemopoietic
RT progenitor cells.";
RL Gene 144:213-219(1994).
DR EMBL: X74861; CA52854.1; -;
DR InterPro: IPR001356; Homeobox.
DR Prodom: PD000010; Homeobox.1.
FT NON_TER 1
FT 39
SQ SEQUENCE 39 AA; 4583 MW; 28BC40613ACE295D CRC64;

Query Match 33.3%; Score 5; DB 4; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESR 13
DB 32 TPESR 36

RESULT 82
ID 055267 PRELIMINARY; PRT; 39 AA.
AC 055267;
DT 01-JUN-1998 (TRENBLREL. 06, Created)
DT 01-JUN-1998 (TRENBLREL. 06, Last sequence update)
DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
DE Tax protein (Fragment).
GN TAX.
OS Human T-cell leukemia virus type II (HTLV-II).
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11909;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-G2;
RA Leon-Ponte M., Echeverria de Perez G., Blanco N., Hengst J., Dube S.,
RA Love J., Polesz B.;
RT "Endemic infection with human T-cell leukemia/lymphoma virus type IIB
RT in Venezuelan Indians: molecular characterization.";
RL Submitted (MAY-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF005396; AAB87754.1; -;
FT NON_TER 1
FT 39
SQ SEQUENCE 39 AA; 4417 MW; 26A991DD18DD65CC CRC64;

Query Match 33.3%; Score 5; DB 15; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HLGPH 6
DB 35 HLGPH 39

RESULT 83
055268

ID 055268 PRELIMINARY; PRT; 39 AA.
AC 055268;
DT 01-JUN-1998 (TRENBLREL. 06, Created)
DT 01-JUN-1998 (TRENBLREL. 06, Last sequence update)
DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
DE Tax protein (Fragment).
GN TAX.
OS Human T-cell leukemia virus type II (HTLV-II).
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11909;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-G4;
RA Leon-Ponte M., Echeverria de Perez G., Blanco N., Hengst J., Dube S.,
RA Love J., Polesz B.;
RT "Endemic infection with human T-cell leukemia/lymphoma virus type IIB
RT in Venezuelan Indians: molecular characterization.";
RL Submitted (MAY-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF005397; AAB87755.1; -;
FT NON_TER 1
FT 39
SQ SEQUENCE 39 AA; 4417 MW; 26A991DD18DD65CC CRC64;

Query Match 33.3%; Score 5; DB 15; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HLGPH 6
DB 35 HLGPH 39

RESULT 84
ID 055269 PRELIMINARY; PRT; 39 AA.
AC 055269;
DT 01-JUN-1998 (TRENBLREL. 06, Created)
DT 01-JUN-1998 (TRENBLREL. 06, Last sequence update)
DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
DE Tax protein (Fragment).
GN TAX.
OS Human T-cell leukemia virus type II (HTLV-II).
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11909;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Y5;
RA Leon-Ponte M., Echeverria de Perez G., Blanco N., Hengst J., Dube S.,
RA Love J., Polesz B.;
RT "Endemic infection with human T-cell leukemia/lymphoma virus type IIB
RT in Venezuelan Indians: molecular characterization.";
RL Submitted (MAY-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF005398; AAB87756.1; -;
FT NON_TER 1
FT 39
SQ SEQUENCE 39 AA; 4417 MW; 26A991DD18DD65CC CRC64;

Query Match 33.3%; Score 5; DB 15; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HLGPH 6
DB 35 HLGPH 39

RESULT 85
ID 055270 PRELIMINARY; PRT; 39 AA.
AC 055270;
DT 01-JUN-1998 (TRENBLREL. 06, Created)
DT 01-JUN-1998 (TRENBLREL. 06, Last sequence update)
DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)

```

DR EMBL: AY049721; AAL06645.1; -.
DR EMBL: AF395588; AAL40694.1; -.
DR InterPro: IPR000345; Cytc_heme_bind.
DR InterPro: IPR000313; PMPW_domain.
DR InterPro: IPR001214; SET.
DR InterPro: IPR001965; ZnF_PHD.
DR Pfam: PF00628; PHD; 3.
DR Pfam: PF00855; PMPW; 2.
DR Pfam: PF00856; SET; 1.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE: PS0280; SET; 1.
KW RECEPTOR_Nuclear_protein.
SQ SEQUENCE 2696 AA; 296648 MW; 4E80E6DCD9A24C81 CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 4; Length 2696;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STEPSR 13
DB 213 STEPSR 218

RESULT 78
ID 030764 PRELIMINARY; PRT; 4340 AA.
AC 030764;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Polypeptide synthase modules 1 and 2.
GN NID1.
OS Streptomyces caelestis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=36816;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL-2821;
RX MEDLINE=98053867; PubMed=9393718;
RA Kakavas S.J., Katz L., Straessl D.;
RT Identification and characterization of the niddamycin polyketide
RT synthase genes from Streptomyces caelestis.
RL J. Bacteriol. 179:7515-7522(1997).
CC -1 SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDS) FAMILY.
DR EMBL: AF016585; AAC46024.1; -.
DR InterPro: IPR002106; AATRNA_ligaseII.
DR InterPro: IPR001227; Ac_transferase.
DR InterPro: IPR002198; ADH_short.
DR InterPro: IPR004410; FADb.
DR InterPro: IPR000794; ketoacyl-synt.
DR InterPro: IPR003880; pantoate_attach.
DR Pfam: PF00698; Acyl_transf; 3.
DR Pfam: PF00106; adh_short; 2.
DR Pfam: PF00109; ketoacyl-synt; 3.
DR Pfam: PF02801; ketoacyl-synt_C; 3.
DR Pfam: PF00550; pp-binding; 3.
DR TIGRfams: TIGR00128; fadp; 3.
DR PROSITE: PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
DR PROSITE: PS00075; AC_DOMAIN; 3.
DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; 2.
DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_1.
KW Oxidoreductase; Phosphopantetheine; transferase.
SQ SEQUENCE 4340 AA; 457589 MW; D59A734CB5FB795D CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 2; Length 4340;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRAA 15
DB 3813 PESRAA 3818

```

```

RESULT 79
ID P70635 PRELIMINARY; PRT; 23 AA.
AC P70635;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE beta2-crystallin (Fragment).
GN CRYB2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=MISTAR;
RX MEDLINE=96305362; PubMed=8706714;
RA Dirks R.P.H., Kraft H.J., Van Genesen S.T., Klok E.J., Pfundt R.,
RA Schoenmakers J.G.G., Lubsen N.H.;
RT "The cooperation between two silencers creates an enhancer element
RT that controls both the lens-preferred and the differentiation stage
RT specific expression of the rat beta2-crystallin gene."
RL Eur. J. Biochem. 239:23-32(1996).
DR EMBL: X83671; CAA58645.1; -.
FT NON_TER 1 1
FT SEQUENCE 23 AA; 2297 MW; 3D8EC25EC44C8C2B CRC64;

Query Match
Best Local Similarity 100.0%; Score 5; DB 11; Length 23;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STEPS 12
DB 12 STEPS 16

RESULT 80
ID 022463 PRELIMINARY; PRT; 26 AA.
AC 022463;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Barperin2 (Fragment).
GN PERM2.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=CV. MOREX;
RA Skadsen R.W., Herbst J.M.;
RT "Barperin2."
RT Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF016328; AAB71681.1; -.
FT NON_TER 1 1
FT SEQUENCE 26 AA; 2569 MW; 4E5598472918A84E CRC64;

Query Match
Best Local Similarity 100.0%; Score 5; DB 10; Length 26;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPHR 7
DB 19 LGPHR 23

RESULT 81

```


RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Massarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhen M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL: AE003530; AAF49616.1; -.
 DR HSP: P05132; IATP.
 DR FLYBASE: FBgn0036511; CG6498.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR001478; PDZ.
 DR InterPro: IPR000961; Pkinase_C.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Pfam: PF00595; PDZ; 1.
 DR Pfam: PF00069; Pkinase; 1.
 DR ProDom: PD000001; Euk_pkinase; 1.
 DR SMART: SM00228; PDZ; 1.
 DR SMART: SM00220; S_TKc; 1.
 DR SMART: SM00133; S_TKc; 1.
 DR PROSITE: PS50106; PDZ; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 2139 AA; 225507 MW; 39FA95E1374A3A CRC64;

Query Match 40.0%; Score 6; DB 5; Length 2139;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPES 12
 Db 1357 RSTPES 1362

RESULT 75
 ID 096RN7 PRELIMINARY; PRT: 2596 AA.
 AC 096RN7;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE NSD1.
 GN NSD1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-21384794; PubMed-11493482;
 RA Jaju R.J., Fidler C., Haas O.A., Strickson A.J., Watkins F., Clark K.,
 RA Cross N.C., Cheng J.F., Aplan P.D., Kearney L., Boulwood J.,
 RA Waltscoat J.S.;
 RT "A novel gene, NSD1, is fused to NUP98 in the t(5;11)(q35;p15.5) in de
 RT novo childhood acute myeloid leukemia.";
 RL Blood 98:1264-1267(2001).
 DR EMBL: AF322907; AA92049.1; -.
 DR InterPro: IPR000345; CytC_heme_bind.
 DR InterPro: IPR000313; PWWP_domain.
 DR InterPro: IPR001214; SET.
 DR InterPro: IPR001965; Znf_PHD.
 DR Pfam: PF00628; PHD; 3.
 DR Pfam: PF00855; PWWP; 1.
 DR Pfam: PF00856; SET; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
 DR PROSITE: PS50280; SET; 1.
 SQ SEQUENCE 2596 AA; 285024 MW; 1CA12DBAD5483379 CRC64;

Query Match 40.0%; Score 6; DB 4; Length 2596;
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 8 STPESR 13
 Db 213 STPESR 218

RESULT 76
 ID 026033 PRELIMINARY; PRT: 2664 AA.
 AC 026033;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Variant-specific surface protein.
 GN VAR-2.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmidium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-FCR3;
 RX MEDLINE-95330813; PubMed-7606788;
 RA Su X.Z., Heatwole V.M., Wertheimer S.P., Guinet F., Herrfeldt J.A.,
 RA Peterson D.S., Ravetch J.A., Wellms T.E.;
 RT "The large diverse gene family var encodes proteins involved in
 RT cytoadherence and antigenic variation of *Plasmodium falciparum*-
 RT infected erythrocytes.";
 RL Cell 82:89-100(1995).
 DR EMBL: I40609; AAA75398.1; -.
 DR InterPro: IPR004258; PFEMP.
 DR Pfam: PF03011; PFEMP; 2.
 SQ SEQUENCE 2664 AA; 302410 MW; 6EA2468511703091 CRC64;

Query Match 40.0%; Score 6; DB 5; Length 2664;
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPES 12
 Db 1335 RSTPES 1340

RESULT 77
 ID 096L73 PRELIMINARY; PRT: 2696 AA.
 AC 096L73;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Androgen receptor associated coregulator 267-b (Putative nuclear
 DE protein NSD1).
 GN ARA267B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-21538797; PubMed-11509567;
 RA Wang X., Yeh S., Wu G., Hsu C.-L., Wang L., Chang T., Yang Y., Guo Y.,
 RA Chang C.;
 RT "Identification and Characterization of a Novel Androgen Receptor
 RT Coregulator ARA267-alpha in Prostate Cancer Cells.";
 RL J. Biol. Chem. 276:40417-40423(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-21590373; PubMed-11733144;
 RA Kurotaki N., Harada N., Yoshitura K., Sugano S., Nakawa N.,
 RA Matsumoto N.;
 RT "Molecular characterization of NSD1, a human homologue of the mouse
 RT Nsd1 gene.";
 RL Gene 279:197-204(2001).

```

DT 01-MAR-2002 (Tremblrel. 20, last annotation update)
DE Similarity to protein kinase.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae:
OC eucots II: Brassicales: Brassicaceae: Arabidopsis.
ON NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RT Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RX MEDLINE-20277480; PubMed-10819329;
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty pl and YAC
RT clones."
RL DNA Res. 7:131-135(2000).
DR EMBL: AB028609; BAB02883.1; -.
DR EMBL: AP000740; BAB02883.1; JOINED.
DR HSSP: P12931; 1PMK.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR004040; STY_pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00069; pkinase; 1
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00221; STYK; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
KM ATP-binding, kinase, transferase.
SQ SEQUENCE 1117 AA; 124861 MW; 799E2D6A7E9750C3 CRC64;

Query Match 40.0%; Score 6; DB 10; Length 1117;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESPAA 15
DB 695 PESRAA 700

RESULT 73
ID 064588 PRELIMINARY; PRT; 1680 AA.
AC 064588;
DT 01-AUG-1998 (Tremblrel. 07, Created)
DT 01-AUG-1998 (Tremblrel. 07, last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, last annotation update)
DE A2934680 protein.
GN A2934680.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots; Rosidae:
OC eucots II: Brassicales: Brassicaceae: Arabidopsis.
ON NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA MEDLINE-20083487; PubMed-10617197;
RA Lin X., Kaul S., Kounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., Vanaken S.E., Unayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Niemman W.C., White O., Eissen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis

```

```

RT thaliana.";
RN Nature 402:761-768(1999).
RL [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA Lin X.;
RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC003096; AAC16266.1; -.
DR InterPro: IPR002198; ADH_short.
DR InterPro: IPR001547; GH_5.
DR InterPro: IPR001611; LRR.
DR Pfam: PF00560; LRR; 4.
DR PROSITE: PS00061; ADH_SHORT; UNKNOWN_1.
DR PROSITE: PS00659; GLYCOSYL_HYDROL_F5; UNKNOWN_1.
SQ SEQUENCE 1680 AA; 183873 MW; E7296D84CA6B53E4 CRC64;

Query Match 40.0%; Score 6; DB 10; Length 1680;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STEPSR 13
DB 248 STEPSR 253

RESULT 74
ID 09VU09 PRELIMINARY; PRT; 2139 AA.
AC 09VU09;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, last annotation update)
DE CG6498 protein.
GN CG6498.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota: Metazoa: Arthropoda: Tracheata: Hexapoda: Insecta;
OC Pterygota: Neoptera: Endopterygota; Diptera: Brachycera; Muscomorpha;
OC Ephydroidea: Drosophilidae; Drosophila.
ON NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortan J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H.C., Blazej R.G., Champs M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abilil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borovaya D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Clawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Dey S., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Galbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Moutulov G., Mlshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauley J.M.,
RA Palzocco M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler J., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

```

DR EMBL: AC108884; AM01139.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 938 AA; 106849 MW; 0CF12969CC9FE88 CRC64;

Query Match 40.0%; Score 6; DB 10; Length 938;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 TPESRA 13
 |||||
 DB 146 TPESRA 151

RESULT 69
 098BM9 PRELIMINARY; PRT; 970 AA.
 AC 098BM9;
 DT 01-OCT-2001 (TReMBLrel. 18, Created)
 DT 01-OCT-2001 (TReMBLrel. 18, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE NAD-dependent formate dehydrogenase alpha subunit.
 GN MLL5396.
 OS Rhizobium loti (Mesorhizobium loti)
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Pyllobacteriaceae; Mesorhizobium.
 NC NCB1_TaxID=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAF30309;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneo T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 Mesorhizobium loti."
 RL DNA Res. 7:331-338(2000).
 DR EMBL: AP003006; BAB51853.1; -.
 DR InterPro: IPR001450; 4Fe4S_ferredoxin.
 DR InterPro: IPR001041; Ferredoxin.
 DR InterPro: IPR001467; Prok_Mboxred.
 DR Pfam: PF00037; fer2; 1.
 DR Pfam: PF00037; fer4; 1.
 DR Pfam: PF00384; Molybdopterin; 1.
 DR Pfam: PF01568; Molybdopterin; 1.
 DR PROSITE: PS00198; 4Fe4S_FERREDOXIN; 2.
 DR PROSITE: PS00551; MOLYBDOTERIN_PROK_1; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 970 AA; 106642 MW; 1E7CAF8FB8AC4E68 CRC64;

Query Match 40.0%; Score 6; DB 16; Length 970;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRA 14
 |||||
 DB 18 TPESRA 23

RESULT 70
 P79749 PRELIMINARY; PRT; 1048 AA.
 AC P79749;
 DT 01-MAY-1997 (TReMBLrel. 03, Created)
 DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Platelet-derived growth factor receptor beta.
 GN PDGFRBETA.
 OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;

OC Tetraodontidae; Takifugu.
 NC NCB1_TaxID=31033;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=97129405; PubMed=8973913;
 RA How G.F., Venkatesh B., Brenner S.;
 RT "Conserved linkage between the puffer fish (Fugu rubripes) and human
 genes for platelet-derived growth factor receptor and macrophage
 RT colony-stimulating factor receptor."
 RL Genome Res. 6:1185-1191(1996).

DR EMBL: U63926; AAC6062.1; -.
 DR HSSP: P11362; IFGK.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR003599; Ig.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00047; Ig; 3.
 DR Pfam: PF00069; pkinase; 1.
 DR ProDom: PD000001; Euk_pkinase; 2.
 DR SMART: SM00409; IG; 3.
 DR SMART: SM00219; TYRKC; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_AMP; 1.
 DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 KW ATP-binding; Receptor; Transferase; Tyrosine-protein kinase.
 SQ SEQUENCE 1048 AA; 117867 MW; 7363FED3BD9B913C CRC64;

Query Match 40.0%; Score 6; DB 13; Length 1048;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLSPH 6
 |||||
 DB 652 SHLSPH 657

RESULT 71
 08RA406 PRELIMINARY; PRT; 1097 AA.
 AC 08RA406;
 DT 01-JUN-2002 (TReMBLrel. 21, Created)
 DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Platelet-derived growth factor receptor beta.
 GN PDGFRB.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NC NCB1_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY;
 RA Wang Y., Cully M.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY090783; AM09098.1; -.
 KW Receptor.
 SQ SEQUENCE 1097 AA; 122827 MW; 5E5540FA0C5CF22B CRC64;

Query Match 40.0%; Score 6; DB 11; Length 1097;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLSPH 6
 |||||
 DB 655 SHLSPH 660

RESULT 72
 09LRV7 PRELIMINARY; PRT; 1117 AA.
 AC 09LRV7;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)

DR Pfam: PF00606; Glycoprotein-B; 1.
 KW Prodom: PD000693; Glycoprot-B; 1.
 KM Signal.
 FT SIGNAL.
 SQ SEQUENCE 865 AA; 98119 MW; 8958E3452EE37D18 CRC64;

Query Match 40.0%; Score 6; DB 12; Length 865;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRA 14
 DB 174 TPESRA 179

RESULT 66

ID Q9RD32 PRELIMINARY; PRT; 888 AA.

AC Q9RD32;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Putative transcriptional regulatory protein.

GN SC00877 OR SC01.10.

OS Streptomyces coelicolor.

OC Bacteria: Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Streptomycetaceae; Streptomyces.

OX NCBI_TaxID=1902;

RP SEQUENCE FROM N.A.

RA STRAIN=A3(2);

RA Saunders D.C., Harris D.;

RL Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.

RP SEQUENCE FROM N.A.

RA STRAIN=A3(2);

RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;

RL Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.

RP SEQUENCE FROM N.A.

RA STRAIN=A3(2);

RA MEDLINE=97000351; PubMed=8843436;

RA Kinsch H., Hopwood D.A.;

RT "A set of ordered cosmids and a detailed genetic and physical map for

the 8 Mb Streptomyces coelicolor A3(2) chromosome.";

RL Mol. Microbiol. 21:77-96(1996).

RP SEQUENCE FROM N.A.

RA STRAIN=A3(2) / M145;

RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,

RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,

RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,

RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,

RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,

RA Radnocsics E., Rajandream M.A., Rutherford K., Rutter S.,

RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,

RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,

RA Hopwood D.A.;

RT "Complete genome sequence of the model actinomycete Streptomyces

coelicolor A3(2).";

RL Nature 417:141-147(2002).

-1- SIMILARITY: BELONGS TO THE LUXR/UHPA FAMILY OF TRANSCRIPTIONAL

REGULATORS.

CC EMBL: AL133422; CAB62668.1; -;

DR InterPro: IPR000792; HTH_LuxR.

DR Pfam: PF00196; Gere; 1.

DR PRINTS: PR00038; HTH_LuxR.

DR Prodom: PD000307; HTH_LuxR; 1.

DR SMART: SM00421; HTH_LuxR; 1.

KW DNA-binding; Transcription regulation.

RL SEQUENCE 888 AA; 91963 MW; 7071FF5261B5B844 CRC64;

Query Match 40.0%; Score 6; DB 16; Length 888;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRA 14
 DB 722 TPESRA 727

RESULT 67

ID Q98280 PRELIMINARY; PRT; 889 AA.

AC Q98280;

DT 01-FEB-1997 (TREMBLrel. 02, Created)

DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE MCV13L.

GN MCV13L.

OS Molluscum contagiosum virus subtype 1 (MCV1).

OC Viruses: dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;

OC Molluscipoxvirus.

OX NCBI_TaxID=10280;

RP SEQUENCE FROM N.A.

RA MEDLINE=96325459; PubMed=8670425;

RA Senkevich T.G., Bugert J.J., Sisler J.R., Koonin E.V., Darai G.,

RA Moss B.;

RT "Genome sequence of a human tumorigenic poxvirus: Prediction of

specific host response-evasion genes.";

RL Science 273:813-816(1996).

RP SEQUENCE FROM N.A.

RA Senkevich T.G., Bugert J.J., Sisler J.R., Koonin E.V., Darai G.,

RA Moss B.;

RL Submitted (JUN-1996) to the EMBL/Genbank/DBJ databases.

DR EMBL: U60315; AAC55241.1; -;

DR InterPro: IPR005058; Pox_P4A.

DR Pfam: PF03395; Pox_P4A; 1.

SQ SEQUENCE 889 AA; 98024 MW; DCDFF181D7EB31FA3 CRC64;

Query Match 40.0%; Score 6; DB 12; Length 889;

Best Local Similarity 100.0%; Pred. No. 1.5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRA 14
 DB 211 TPESRA 216

RESULT 68

ID Q85SP5 PRELIMINARY; PRT; 938 AA.

AC Q85SP5;

DT 01-JUN-2002 (TREMBLrel. 21, Created)

DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Hypothetical 106.8 kDa protein.

GN OSUBM0058B20.21

OS Oryza sativa (japonica cultivar-group).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzae; Oryza.

OX NCBI_TaxID=39947;

RP SEQUENCE FROM N.A.

RA STRAIN=CV. NIPPONBARE;

RA Kuc K., Nascimeto L., Zutavern T., Balija V., Bell M., Baker J.,

RA Santos L., Miller B., Katzenberger F., Muller S., King L., Yang C.,

RA O'Shaughnessy A., Palmer L., Dedhia N.;

RT "Genomic sequence for Oryza sativa, Nipponbare strain, clone

OSUNB0058B20, from chromosome 10, complete sequence.";

RL Submitted (APR-2002) to the EMBL/Genbank/DBJ databases.

RT transcripts from full-length clone of arlichoke mottle crinkle
RT virus.";
RL J. Gen. Virol. 75:1515-1524(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-AMCV-BARI DR.GALLITELLI ISOLATE;
RA TAVAZZA M.;
RL Submitted (OCT-1991) to the EMBL/GenBank/DBJ databases.
DR EMBL: X62493; CAA44356.1; -.
DR InterPro: IPR002564; PV_RdRp.
DR Pfam: PF01615; PV_RdRp.1
SQ SEQUENCE 818 AA; 92157 MW; 247440E0A692280 CRC64;

Query Match 40.0%; Score 6; DB 12; Length 818;
Best Local Similarity 100.0%; Pred. No.1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 TPESRA 14
Db 767 TPESRA 772

RESULT 62
Q9WJD6 PRELIMINARY; PRT; 818 AA.
ID Q9WJD6
AC Q9WJD6
DT 01-NOV-1999 (TREMblrel. 12, Created)
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE Translated region.
OS Cymbidium ringspot virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tombusviridae;
OC Tombusvirus.
OX NCBI_TaxID=12144;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8936663; PubMed=2771646;
RA Grieco F., Burgyan J., Russo M.;
RT "The Nucleotide sequence of Cymbidium ringspot virus RNA.";
RL Nucleic Acids Res. 17:6383-6383(1989).
DR EMBL: X15511; CAB36439.1; -.
DR InterPro: IPR002564; PV_RdRp.
DR Pfam: PF01615; PV_RdRp.1.
SQ SEQUENCE 818 AA; 91946 MW; DC4E84949410F42D CRC64;

Query Match 40.0%; Score 6; DB 12; Length 818;
Best Local Similarity 100.0%; Pred. No.1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 TPESRA 14
Db 767 TPESRA 772

RESULT 63
Q66101 PRELIMINARY; PRT; 851 AA.
ID Q66101
AC Q66101
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Replicase.
GN ORF2.
OS Carnation Italian ringspot virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tombusviridae;
OC Tombusvirus.
OX NCBI_TaxID=39443;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96105357; PubMed=7503699;
RA Rubino L., Burgyan J., Russo M.;
RT "Molecular cloning and complete nucleotide sequence of carnation
Italian ringspot tombusvirus genomic and defective interfering RNAs.";

RL Arch. Virol. 140:2027-2039(1995).
DR EMBL: X85215; CAA59478.1; -.
DR InterPro: IPR002564; PV_RdRp.
DR Pfam: PF01615; PV_RdRp.1
SQ SEQUENCE 851 AA; 95389 MW; 0A6115E2A817E0B CRC64;

Query Match 40.0%; Score 6; DB 12; Length 851;
Best Local Similarity 100.0%; Pred. No.1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 TPESRA 14
Db 800 TPESRA 805

RESULT 64
Q9NO53 PRELIMINARY; PRT; 858 AA.
ID Q9NO53
AC Q9NO53
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Zinc finger protein Cezanne.
GN CEZANNE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21356309; PubMed=11463333;
RA Evans P.C., Taylor E.R., Coadwell J., Heynink K., Beyaert R.,
RA Kishaw P.J.;
RT "Isolation and characterization of two novel A20-like proteins.";
RL Biochem. J. 357:617-623(2001).
DR EMBL: AJ293573; CAB97494.1; -.
DR InterPro: IPR003323; OTU.
DR InterPro: IPR002653; ZnF_A20.
DR SMART: SM00259; ZnF_A20; 1.
DR PROSITE: PSS0802; OTU: 1.
SQ SEQUENCE 858 AA; 94401 MW; A7D6B1D280C9387F CRC64;

Query Match 40.0%; Score 6; DB 4; Length 858;
Best Local Similarity 100.0%; Pred. No.1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPES 12
Db 474 RSTPES 479

RESULT 65
Q83291 PRELIMINARY; PRT; 865 AA.
ID Q83291
AC Q83291
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Glycoprotein 100 precursor.
GN GB.
OS Marek disease virus type 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=38013;
RN [1]
RP SEQUENCE OF 1-498 FROM N.A.
RC STRAIN-JM;
RA Souslopavov M.A., Bakhtina M.M., Krendelshtchikov A.V., Babkin I.V.;
RT "PCR-mediated cloning and sequencing of gene encoding the glycoprotein
complex gp100, gp60, gp48 (b-antigen).";
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: X91985; CAA63039.1; -.
DR InterPro: IPR000234; Glycoprot_B.

QY 9 TPESRA 14
|||||
DB 174 TPESRA 179

RESULT 58

094GT7 PRELIMINARY; PRT; 808 AA.
AC 094GT7;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Putative retrotransposable elements TMP2.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
OX NCBL_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Wing R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,
RT Sasaki C., Henry D., Oates R., Simmons J.;
RL "Rice Genomic Sequence."
DR EMBL: AC087723; AAK95676.1; -;
DR InterPro: IPR004242; Transposase_21.
DR Pfam: PF02992; Transposase_21; 1.
SQ SEQUENCE 808 AA; 93019 MW; 87EAD18974FB3F66 CRC64;

Query Match

Best Local Similarity 40.0%; Score 6; DB 10; Length 808;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 TPESRA 14
|||||
DB 446 TPESRA 451

RESULT 59

088485 PRELIMINARY; PRT; 818 AA.
AC 088485;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE P92 protein.
OS Tomato bushy stunt virus (Strain Cherry) (TBSV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tombusviridae;
OC Tombusvirus.
OX NCBL_TaxID=12147;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHERRY;
RA MEDLINE=89163266; PubMed=2922927;
RA Hillman B.I., Hearne P., Rochon D., Morris T.J.;
RT "Organization of tomato bushy stunt virus genome: characterization of
RT the coat protein gene and the 3' terminus."
RL Virology 169:42-50(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CHERRY;
RX MEDLINE=90281577; PubMed=2353450;
RA Hearne P.Q., Knorr D.A., Hillman B.I., Morris T.J.;
RT "The complete genome structure and synthesis of infectious RNA from
RT clones of tomato bushy stunt virus."
RL Virology 177:141-151(1990).
DR EMBL: M21958; AAB02535.1; -;
DR InterPro: IPR002564; PV_RdRp.
DR Pfam: PF01615; PV_RdRp; 1.
SQ SEQUENCE 818 AA; 92251 MW; D61AD04D5F94BC42 CRC64;

Query Match

40.0%; Score 6; DB 12; Length 818;

Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 TPESRA 14
|||||
DB 767 TPESRA 772

RESULT 60

P89211 PRELIMINARY; PRT; 818 AA.
AC P89211;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE 92k protein.
OS Tomato bushy stunt virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tombusviridae;
OC Tombusvirus.
OX NCBL_TaxID=12145;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEPPER;
RX MEDLINE=98285798; PubMed=9621100;
RA Havelde Z., Szilva G., Burgan J.;
RT "Characterization of the molecular mechanism of defective interfering
RT RNA-mediated symptom attenuation in tombusvirus-infected plants."
RL J. Virol. 72:6251-6256(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=PEPPER;
RX MEDLINE=20473138; PubMed=11018282;
RA Szilva G., Salamon P., Burgan J.;
RT "The complete nucleotide sequence and synthesis of infectious RNA of
RT genomic and defective interfering RNAs of TBSV-P."
RL Virus Res. 69:131-136(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=PEPPER;
RA Szilva G., Burgan J.;
RT Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U80935; AAC32730.1; -;
DR InterPro: IPR002564; PV_RdRp.
DR Pfam: PF01615; PV_RdRp; 1.
SQ SEQUENCE 818 AA; 92566 MW; 53C96D56E927CBB CRC64;

Query Match

Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 TPESRA 14
|||||
DB 767 TPESRA 772

RESULT 61

096606 PRELIMINARY; PRT; 818 AA.
AC 096606;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Protein of 92 kDa.
OS Artichoke mottled crinkle virus (AMCV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tombusviridae;
OC Tombusvirus.
OX NCBL_TaxID=12142;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AMCV-BART DR. GALLITELLI ISOLATE;
RX MEDLINE=94292901; PubMed=8021582;
RA Tavaza M., Lucifora A., Calogero A., Pay A., Tavaza R.;
RT "Nucleotide sequence, genomic organization and synthesis of infection

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=SPLEEN;
 RA Ohara O., Nagase T., Kikuno R., Okumura K.;
 RT "The nucleotide sequence of a long CDNA clone isolated from human spleen."
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK074045; BAB84871.1; -.
 FT NON_TER
 SQ SEQUENCE 678 AA; 70813 MW; 4847E47E56F5D696 CRC64;

Query Match 40.0%; Score 6; DB 4; Length 678;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 SPESR 13
 DB 10 SPESR 15

RESULT 54
 O80GX3 PRELIMINARY; PRT; 735 AA.
 AC O80GX3;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE XEXT1.
 GN XEXT1.
 OS Xenopus laevis (African clawed frog).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 CC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8335;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Katada T., Oogami S., Shima N., Kinoshita T.;
 RT "cDNA cloning and distribution of the XEXT1, Xenopus homologue of
 RL Ext1."
 RL Dev. Genes Evol. 0:0-0(2002).
 DR EMBL; AB048247; BAB87180.1; -.
 SQ SEQUENCE 735 AA; 84337 MW; 5DDA8010EF2352A1 CRC64;

Query Match 40.0%; Score 6; DB 13; Length 735;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRAA 15
 DB 39 PESRAA 44

RESULT 55
 O9E352 PRELIMINARY; PRT; 787 AA.
 ID O9E352;
 AC O9E352;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE P92 protein.
 OS Maize necrotic streak virus.
 CC Viruses; ssRNA positive-strand viruses, no DNA stage; Tombusviridae.
 OX NCBI_TaxID=137556;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Louie R., Redinbaugh M.G., Gordon D.T., Abt J.J., Anderson R.J.;
 RT "Maize Necrotic Streak Virus, a New Maize Virus with Similarity to
 RL Species of the Family Tombusviridae."
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF266518; AAG21222.1; -.

DR InterPro: IPR002564; PV_RdRp.
 DR Pfam: PF01615; PV_RdRp; 1.
 SQ SEQUENCE 787 AA; 88629 MW; 92202B2F5A7C0BEA CRC64;

Query Match 40.0%; Score 6; DB 12; Length 787;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRA 14
 DB 736 TPESRA 741

RESULT 56
 O98Y45 PRELIMINARY; PRT; 805 AA.
 ID O98Y45;
 AC O98Y45;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Glycoprotein B (Fragment).
 OS Turkey herpesvirus.
 CC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 CC Alphaherpesvirinae; Marek's disease-like viruses.
 OX NCBI_TaxID=10390;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Davidson I.;
 RT "Marek's disease virus - chicken isolate, gb gene."
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF350325; AAK37556.2; -.
 DR InterPro: IPR000234; Glycoprot_B.
 DR Pfam: PF00606; Glycoprotein_B; 1.
 DR ProDom: PD000693; Glycoprot_B; 1.
 FT NON_TER
 SQ SEQUENCE 805 AA; 91111 MW; F88B4410A85E15AA CRC64;

Query Match 40.0%; Score 6; DB 12; Length 805;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRA 14
 DB 174 TPESRA 179

RESULT 57
 O98Y44 PRELIMINARY; PRT; 805 AA.
 ID O98Y44;
 AC O98Y44;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Glycoprotein B (Fragment).
 OS Turkey herpesvirus.
 CC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 CC Alphaherpesvirinae; Marek's disease-like viruses.
 OX NCBI_TaxID=10390;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Davidson I.;
 RT "Marek's disease virus - turkey isolate, gb."
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF350326; AAK37557.2; -.
 DR InterPro: IPR000234; Glycoprot_B.
 DR Pfam: PF00606; Glycoprotein_B; 1.
 DR ProDom: PD000693; Glycoprot_B; 1.
 FT NON_TER
 SQ SEQUENCE 805 AA; 91115 MW; FC8EB090885E15AA CRC64;

Query Match 40.0%; Score 6; DB 12; Length 805;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DR ProDom; PD000001; Euk_pkinase; 1.
 DR SMART; SM00370; LRR; 5.
 DR SMART; SM00220; S_TKc; 1.
 DR SMART; SM00219; TyrKc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR APP-binding; Kinase; transferase.
 SO SEQUENCE 614 AA; 67963 MW; C61DAD9A8B3A6B74 CRC64;

Query Match 40.0%; Score 6; DB 10; Length 614;
 Best Local Similarity 100.0%; Pred. No. 1.le+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 RSTPES 12
 Db 516 RSTPES 521

RESULT 50

ID 068077 PRELIMINARY; PRT; 617 AA.
 AC 068077;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE Hypothetical 66.1 kDa protein.
 OS Rhodobacter capsulatus (Rhodospseudomonas capsulata).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
 OC Rhodobacter.
 NCBI_TaxID=1061;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SB1003;
 RX MEDLINE-97404404; PubMed-9256491;
 RA Vitek C., Paces V., Maltsev N., Paces J., Haselkorn R., Fomstein M.;
 RT "Sequence of a 189-kb segment of the chromosome of Rhodobacter
 capsulatus SB1003."
 RL Proc. Natl. Acad. Sci. U.S.A. 94:9384-9388(1997).
 DR EMBL; AF010496; AAC16163.1; -.
 DR InterPro; IPR001440; TPR.
 DR Pfam; PF00515; TPR; 2.
 KW Hypothetical protein.
 SO SEQUENCE 617 AA; 66073 MW; 32E77E32579EE417 CRC64;

Query Match 40.0%; Score 6; DB 2; Length 617;
 Best Local Similarity 100.0%; Pred. No. 1.le+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 PESRAA 15
 Db 470 PESRAA 475

RESULT 51

ID 08TUB4 PRELIMINARY; PRT; 637 AA.
 AC 08TUB4;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Cleavage and polyadenylation specificity factor.
 GN MA3874.
 OS Methanosarcina acetivorans.
 OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
 OC Methanosarcinaceae; Methanosarcina.
 NCBI_TaxID=2214;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C2A / ATCC 35395 / DSM 2834;
 RX MEDLINE-21929760; PubMed-11932238;
 RA Galagan J.E., Nusbaum C., Roy A., Endlitz M.G., Macdonald P.,
 RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Alnoor D., Brown A.,
 RA Allen N., Naylor J., Stange-Thomann N., Dearellano K., Johnson R.,

RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
 RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
 RA Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,
 RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
 RA Springer T.A., Umeyan L.A., White O., White R.H., de Macario E.C.,
 RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
 RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
 RA Metcalf W.W., Birren B.;
 RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
 and physiological diversity."
 RL Genome Res. 12:532-542(2002).
 DR EMBL; AE011098; AAM07225.1; -.
 KW Complete proteome.
 SO SEQUENCE 637 AA; 71847 MW; 7A8F536DD9B9E5FD CRC64;

Query Match 40.0%; Score 6; DB 17; Length 637;
 Best Local Similarity 100.0%; Pred. No. 1.le+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 STPESR 13
 Db 200 STPESR 205

RESULT 52

ID 08YY1 PRELIMINARY; PRT; 640 AA.
 AC 08YY1;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical protein Al10711.
 GN Al10711.
 OS Anabaena sp. (strain PCC 7120).
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 NCBI_TaxID=103690;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-21595285; PubMed-11759840;
 RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
 RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
 RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
 RA Yasuda M., Tabata S.;
 RT "Complete genomic sequence of the filamentous nitrogen-fixing
 cyanobacterium Anabaena sp. strain PCC 7120."
 RL DNA Res. 8:205-213(2001).
 DR EMBL; AP003583; BAB72668.1; -.
 DR InterPro; IPR001023; Hsp70.
 DR ProDom; PD000089; Hsp70; 1.
 DR PROSITE; PS00329; HSP70_2; UNKNOWN_1.
 KW Hypothetical protein; Complete proteome.
 SO SEQUENCE 640 AA; 72075 MW; FFD1160A2E98232 CRC64;

Query Match 40.0%; Score 6; DB 16; Length 640;
 Best Local Similarity 100.0%; Pred. No. 1.le+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 PESRAA 15
 Db 142 PESRAA 147

RESULT 53

ID 08TES7 PRELIMINARY; PRT; 678 AA.
 AC 08TES7;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE FLJ00103 protein (Fragment).
 GN FLJ00103.
 OS Homo sapiens (Human).

Query Match
Best Local Similarity 40.0%; Score 6; DB 16; Length 429;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRAA 15
Db 171 PESRAA 176

RESULT 44
O9SR06 PRELIMINARY; PRT; 444 AA.

AC O9SR06: 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE F7018.15 protein (Hypothetical 48.4 kDa protein).
GN F7018.15 OR AT3G04680, F7018.15.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,
RA Romling C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome III BAC F7018 genomic sequence."
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Ban J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC011437; AAF04897.1; -;
DR EMBL: AT062866; ALU32944.1; -;
KW Hypothetical protein.
SQ SEQUENCE 444 AA; 48421 MW; 04968918D3A539AA CRC64;

Query Match
Best Local Similarity 40.0%; Score 6; DB 10; Length 444;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRAA 15
Db 235 PESRAA 240

RESULT 45
O95007 PRELIMINARY; PRT; 468 AA.

AC O95007: 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Y66D12A.4 protein.
GN Y66D12A.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Sulston J.E.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.

RX MEDLINE-99069613; PubMed-9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL: AL161712; CAC70130.1; -;
SQ SEQUENCE 468 AA; 54341 MW; 694F9ED4EA7375C8 CRC64;

Query Match
Best Local Similarity 40.0%; Score 6; DB 5; Length 468;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 GPHRST 9
Db 309 GPHRST 314

RESULT 46
O925F7 PRELIMINARY; PRT; 563 AA.

AC O925F7: 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Platelet-derived growth factor receptor beta (fragment).
GN PDGFRB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WISTAR;
RA Okuyama H., Shimahara Y., Kawada N., Yamaoka Y.;
RT "Regulation of cell growth by redox-mediated extracellular proteolysis
RT of platelet-derived growth factor receptor beta."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF359356; AAK43716.1; -;
DR InterPro: IPR007119; Euk_pkinase.
DR InterPro: IPR001824; RTkinaseII.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF000069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 2.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; UNKNOWN_1.
DR PROSITE: PS00240; RECEPTOR_TYR_KIN_III; UNKNOWN_1.
KW Receptor.
FT NON_TER 1
FT NON_TER 563
SQ SEQUENCE 563 AA; 62986 MW; 80D2CFD256615285 CRC64;

Query Match
Best Local Similarity 40.0%; Score 6; DB 11; Length 563;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGGH 6
Db 128 SHLGGH 133

RESULT 47
O9RJ02 PRELIMINARY; PRT; 575 AA.

AC O9RJ02: 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Putative sensor kinase.
GN SC00588 OR SCF55.12C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycetaceae; Streptomyces.

```

RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL: AE005712; AKK22376.1; -.
DR TIGR: CC0389; -.
DR InterPro: IPR000195; RabGAP_TBC.
DR InterPro: IPR001440; TPR.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 408 AA; 44585 MW; B6129BB3D57E507 CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 16; Length 408;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRA 15
DB 117 PESRA 122

RESULT 41
OBY1M1 PRELIMINARY; PRT; 413 AA.
AC OBY1M1;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)
DE Probable phosphoribosylaminoimidazole carboxylase ATPase subunit
DE protein (EC 4.1.1.21).
GN PURK OR RSC0576 OR RS04889.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-GMI1000;
RX MEDLINE-21681879; PubMed-11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Artlet M., Billault A., Brotlier P., Camus J.C., Cattolico L.,
RA Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Sigler P., Thebault P., Whalen M., Winkler P., Levy M.,
RA Weisenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL: AL646060; CAD14106.1; -.
DR InterPro: IPR003135; ATP-grasp.
DR Pfam: PF02222; ATP-grasp; 1.
DR TIGRFAMS: TIGR01161; purK; 1.
KW Lyase; Complete proteome.
SQ SEQUENCE 413 AA; 44581 MW; F20A6E40055DEA0C CRC64;

Query Match
Best Local Similarity 40.0%; Score 6; DB 16; Length 413;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 PESRA 14
DB 14 PESRA 19

RESULT 42
O99XK9 PRELIMINARY; PRT; 426 AA.
AC O99XK9;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)
DE Putative histidine-tRNA ligase (EC 6.1.1.21).
GN HISS OR SPY2157.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SF370 / ATCC 700294 / SEROTYPE M1;
RX MEDLINE-21192684; PubMed-11296296;
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
DR EMBL: AE006634; AKK34794.1; -.
DR HSSP: O32422; LOBO.
DR InterPro: IPR002106; AtrRNA_ligaseII.
DR InterPro: IPR004154; HGTP_antlicodon.
DR InterPro: IPR004516; HISS.
DR InterPro: IPR002314; tRNA-synt_2b.
DR Pfam: PF03129; HGTP_antlicodon; 1.
DR Pfam: PF00587; tRNA-synt_2b; 1.
DR TIGRFAMS: TIGR00442; hiss; 1.
DR PROSITE: PS00339; AA_TRNA_LIGASE_IL2; UNKNOWN_1.
KW Ligase; Complete proteome.
SQ SEQUENCE 426 AA; 48200 MW; 3E0E3B3FE9E948B CRC64;

```

```

Query Match
Best Local Similarity 100.0%; Score 6; DB 16; Length 426;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRA 15
DB 171 PESRA 176

RESULT 43
O97NC9 PRELIMINARY; PRT; 429 AA.
AC O97NC9;
DT 01-OCT-2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)
DE Histidyl-tRNA synthetase.
GN SP2121.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TIGR4;
RX MEDLINE-21357209; PubMed-11463916;
RA Tettelein H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., Debey R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwin O., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae.";
RL Science 293:498-506(2001).
DR EMBL: AE007501; AKK76180.1; -.
DR TIGR: SP2121; -.
DR InterPro: IPR002106; AtrRNA_ligaseII.
DR InterPro: IPR004154; HGTP_antlicodon.
DR InterPro: IPR004516; HISS.
DR InterPro: IPR002314; tRNA-synt_2b.
DR Pfam: PF03129; HGTP_antlicodon; 1.
DR Pfam: PF00587; tRNA-synt_2b; 1.
DR TIGRFAMS: TIGR00442; hiss; 1.
DR PROSITE: PS00339; AA_TRNA_LIGASE_IL2; UNKNOWN_1.
KW Aminoacyl-tRNA synthetase; Complete proteome.
SQ SEQUENCE 429 AA; 48654 MW; E0ACAB4FBFA3B10 CRC64;

```

SQ SEQUENCE 389 AA; 42559 MW; C8E6F38E21CE7B5F CRC64;
 Query Match 40.0%; Score 6; DB 3; Length 389;
 Best Local Similarity 100.0%; Pred. No. 74;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 10 PESRAA 15
 Db 17 PESRAA 22
 RESULT 38
 ID 007569 PRELIMINARY; PRT; 401 AA.
 AC 007569;
 DT 01-JUN-1997 (TReMBLrel. 04, Created)
 DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Hypothetical protein yhjo.
 GN yho.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Bacillaceae; Bacillus.
 NC NCBL_TaxID=1423;
 RX MEDLINE-98044033; PubMed-9384377;
 RC STRAIN-168;
 RN Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 Borriess R., Brunsler L., Brans A., Braun M., Brignell S.C., Bron S.,
 Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
 Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
 Denicot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
 Ertlan K.D., Errington J., Fabret C., Ferrati E., Foulger D.,
 Fultz C., Fujita M., Fujita Y., Funo S., Galizzi A., Galleron N.,
 Gilm S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
 Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
 Joris B., Karamata D., Kasahara Y., Klaerr-Bianchard M., Klein C.,
 Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
 Medina N., Mellado R.P., Mizuno W., Moestl D., Nakai S., Noback M.,
 Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,
 Paro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadaie Y.,
 Sato T., Scanlan E., Schleich S., Schroeter R., Scifone F.,
 Sekiuchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
 Sorokin A., Taccoti E., Takagi T., Takahashi H., Takemaru K.,
 Takeuchi M., Takemoshi A., Tanaka T., Terpstra P., Tognoni A.,
 Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
 Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
 RT "The complete genome sequence of the gram-positive bacterium *Bacillus*
 subtilis.";
 RN Nature 390:249-256(1997).
 RP SEQUENCE FROM N.A.
 RC STRAIN-168;
 RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.,
 Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 DR EMBL: Y14081; CAA744477.1; -;
 DR EMBL: Z99109; CAB12898.1; -;

DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF00083; sugar_tr.1.
 KW Hypothetical protein; Transmembrane; Complete proteome.
 SQ SEQUENCE 401 AA; 44065 MW; 24F6863D6846BF9 CRC64;
 Query Match 40.0%; Score 6; DB 16; Length 401;
 Best Local Similarity 100.0%; Pred. No. 75;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 8 STPESR 13
 Db 131 STPESR 136
 RESULT 39
 ID 090FF4 PRELIMINARY; PRT; 407 AA.
 AC 090FF4;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE Hypothetical 43.8 kDa protein (fragment).
 GN DKFZP434H0717.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NC NCBL_TaxID=9606;
 RX [1]
 RC TISSUE-TESTIS;
 RA Bloecker H., Boecher M., Brandt P., Mewes H.W., Gassenhuber J.,
 Wleemann S.;
 RN Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL122102; CAB59268.1; -;
 DR InterPro: IPR002653; ZnF_A20.
 DR SMART: SM00259; ZnF_A20; 1.
 FT NON_TER 1
 SQ SEQUENCE 407 AA; 43800 MW; FDF4F3CE2AC0D83C CRC64;
 Query Match 40.0%; Score 6; DB 4; Length 407;
 Best Local Similarity 100.0%; Pred. No. 76;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 7 RSTPES 12
 Db 23 RSTPES 28
 RESULT 40
 ID 09AB45 PRELIMINARY; PRT; 408 AA.
 AC 09AB45;
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Hypothetical protein CC0389.
 GN CC0389.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
 OC Caulobacter.
 NC NCBL_TaxID=155892;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 19089 / CB15;
 RX MEDLINE-21173698; PubMed-11259647;
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 Potocka I., Nelson W.C., Newton A., Stephens C., Phadt N.D., Ely B.,
 Debay R.T., Dodson R.J., Durkin A.S., Gysin M.L., Haft D.H.,
 Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
 Uteback F., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
 Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;

```

SQ  SEQUENCE 301 AA; 32199 MW; D50D1363C6BEC892 CRC64;
Query Match 40.0%; Score 6; DB 16; Length 301;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRAA 15
    |||||
DB 13 PESRAA 18

RESULT 34
ID 08S6J9 PRELIMINARY; PRT; 315 AA.
AC 08S6J9;
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DE Hypothetical 36.4 kDa protein.
GN OSJNBA0019N10.27.
OS Oryza sativa (Rice).
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
OC Spermatophyta: Magnoliophyta: Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_Taxid=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIPPONBARE;
RA McCombie W.R., de la Bastide M., Spiegel L., Preston R., Kirchoff K.,
RA Kuit K., Nascimento L., Zutavern T., Ballja V., Bell M., Baker J.,
RA Santos L., Miller B., Katzenberger F., Muller S., King L., Yang C.,
RA O'Shaughnessy A., Palmer L., Dedola N.;
RT "Genomic sequence for Oryza sativa, Nipponbare strain, clone
RT OSJNBA0019N10, from chromosome 10, complete sequence.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC092748; AAM01089.1; -.
KW Hypothetical protein.
SQ SEQUENCE 315 AA; 36447 MW; D09A13D94EA7EDB1 CRC64;

Query Match 40.0%; Score 6; DB 10; Length 315;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRA 14
    |||||
DB 118 TPESRA 123

RESULT 35
ID 09XEP5 PRELIMINARY; PRT; 343 AA.
AC 09XEP5;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE Hypothetical 38.5 kDa protein.
OS Sorghum bicolor (Sorghum) (Sorghum vulgare).
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;
OC Spermatophyta: Magnoliophyta: Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Sorghum.
OX NCBI_Taxid=4558;
RN [1]
RP SEQUENCE FROM N.A.
RA Llaica V., Lou A., Young S., Messing J.;
RT "Retriable elements of Sorghum bicolor.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF114171; AAD27563.1; -.
KW Hypothetical protein.
SQ SEQUENCE 343 AA; 38484 MW; 26FE636599BA8C1 CRC64;

Query Match 40.0%; Score 6; DB 10; Length 343;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRAA 15
    |||||
DB 123 PESRAA 128

RESULT 36
ID 091T06 PRELIMINARY; PRT; 365 AA.
AC 091T06;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 197.
OS Tupala herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae.
OX NCBI_Taxid=10397;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2;
RX MEDLINE=21211637; PubMed=11312357;
RA Bahr U., Darai G.;
RT "Analysis and characterization of the Complete Genome of Tupala (Tree
RT Shrew) Herpesvirus.";
RL J. Virol. 75:4854-4870(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=2;
RA Darai G., Bahr U.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF281817; AAK57081.1; -.
SQ SEQUENCE 365 AA; 41263 MW; F97E5B06AA77267 CRC64;

Query Match 40.0%; Score 6; DB 12; Length 365;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRAA 15
    |||||
DB 258 PESRAA 263

RESULT 37
ID 09HEP7 PRELIMINARY; PRT; 369 AA.
AC 09HEP7;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE PKA regulatory subunit.
OS Blumeria graminis.
OC Eukaryota: Fungi; Ascomycota; Perizomycotina; Leotiomycetes;
OC Erysiphales; Erysiphaceae; Blumeria.
OX NCBI_Taxid=34373;
RN [1]
RP SEQUENCE FROM N.A.
RA Blidslev L.;
RL Thesis (2001), Department of Molecular Biology,
RL University of Copenhagen, Copenhagen, Denmark.
DR EMBL: AJ304829; CAC19660.1; -.
DR HSSP: P00515; 2APK.
DR InterPro: IPR000515; BPD_transp.
DR InterPro: IPR002373; CNMP_kin.
DR Pfam: PF00027; CNMP_binding.
DR PRINTS: PR00103; CAMPKINASE.
DR SMART: SM00100; CNMP; 2.
DR PROSITE: PS00402; BPD_TRANSF_INN_MEMBER; UNKNOWN_1.
DR PROSITE: PS00888; CNMP_BINDING_1; 1.
DR PROSITE: PS00889; CNMP_BINDING_2; 2.
DR PROSITE: PS00402; CNMP_BINDING_3; 2.
```

```

09BDK8
ID 09BDK8 PRELIMINARY: PRT; 259 AA.
AC 09BDK8;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Platelet-derived growth factor receptor beta (Fragment).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Remillard P.E., Lacroix D.A., Murphy B.D.;
RA Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.
CC -1 SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -1 SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
DR EMBL; AF347051; AAK31152.1; -.
DR HSSP; P11362; 1FGK.
DR InterPro: IPR000719; Euk.pkinase.
DR InterPro: IPR001824; RTKinaseIII.
DR InterPro: IPR002290; Ser_Thr_pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk.pkinase; 2.
DR SMART; SM00220; S.TKc; 1.
DR SMART; SM00219; TYRc; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
KW Glycoprotein; Immunoglobulin domain; Phosphorylation; Receptor;
KW Transmembrane.
FT NON_TER 1 1
FT NON_TER 259 259
SQ SEQUENCE 259 AA; 28776 MW; F1A432566282D951 CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 6; Length 259;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHLGGP 6
Db 5 SHLGGP 10

RESULT 32
09A6S2
ID 09A6S2 PRELIMINARY: PRT; 262 AA.
AC 09A6S2;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Exodeoxyribonuclease III.
OS CC2011.
OC Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-21173698; PubMed=11259647;
RC STRAIN-ATCC 19089 / CB15;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eissen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwin M.L., Haft D.H.,
RA Kolony J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Uterback T., Tran K., Wolf A., Vamathavan J., Ernoiaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RA "Complete genome sequence of Caulobacter crescentus."
RA Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AE005874; AAK23986.1; -.

```

```

DR HSSP; P09030; IAKO.
DR TIGR; CC2011; -.
DR InterPro: IPR000097; Appendocnclsel.
DR InterPro: IPR004442; EXODNase_III.
DR InterPro: IPR004808; EXOIII_xth.
DR InterPro: IPR005135; Exo_endo_phos.
DR Pfam; PF03372; Exo_endo_phos; 1.
DR TIGRFAMs; TIGR00195; exodnase_III; 1.
DR TIGRFAMs; TIGR00633; xth; 1.
DR PROSITE; PS00728; AP_NUCLEASE_FL_3; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 262 AA; 29351 MW; DA57A513BC7E415D CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 16; Length 262;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 PESRAA 15
Db 176 PESRAA 181

RESULT 33
09RI46
ID 09RI46 PRELIMINARY: PRT; 301 AA.
AC 09RI46;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Hypothetical protein SC00200.
GN SC00200 OR SCJ12.12C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-A3(2);
RC STRAIN-A3(2);
RA Thompson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
RA Submitted (AUG-1999) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN-A3(2);
RC STRAIN-A3(2);
RA Redenbach M., Kieser H.M., Denapalce D., Elchner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RA "A set of ordered cosmid and a detailed genetic and physical map for
RA the 8 Mb Streptomyces coelicolor A3(2) chromosome."
RA Mol. Microbiol. 21:77-96(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Croft A., Fraser A., Goble A., Haldago J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RA "Complete genome sequence of the model actinomycete Streptomyces
RA coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL109989; CAB53424.1; -.
DR InterPro: IPR000041; USP.
DR Pfam; PF00582; USP; 2.
KW Hypothetical protein.

```

```

Query Match          40.0%; Score 6; DB 12; Length 243;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 PHSRP 10
    |||||
Db 221 PHSRP 226

RESULT 28
OY0G6 PRELIMINARY; PRT; 255 AA.
AC O8Y0G6;
DT 01-MAR-2002 (TREMblrel. 20, Created)
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Probable transcriptional regulatory DNA-binding transcription
DE regulator protein.
GN RSC1078 OR RS04109.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia
CX NCBI_TaxID=305;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-GM1100;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Ariat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandel N., Choise N., Claudel-Renard C., Cunac S., Demange N.,
RA Gaupier C., Lavié M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Siguler P., Thebaud P., Whalen M., Winkler P., Lévy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL, AL646062; CAD14780.1; -
DR InterPro; IPR000524; HTH_GNTR.
DR Pfam; PF00392; gntr. 1.
DR PRINTS; PR00035; HTHGNTR.
DR SMART; SM00345; HTH_GNTR; 1.
DR PROSITE; PS00043; HTH_GNTR_FAMILY; 1.
KW Complete proteome.
SO SEQUENCE 255 AA; 27440 MW; 3DB0775927E4DE15 CRC64;

Query Match          40.0%; Score 6; DB 16; Length 255;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRA 14
    |||||
Db 194 TPESRA 199

RESULT 29
OY8J74 PRELIMINARY; PRT; 256 AA.
ID O98J74;
AC O98J74;
DT 01-OCT-2001 (TREMblrel. 18, Created)
DT 01-OCT-2001 (TREMblrel. 18, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Probable exonuclease.
GN ML2070.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
CX NCBI_TaxID=381;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-MAFF30309;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,

```

```

RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP002998; BAB49292.1; -
DR InterPro; IPR000097; Apendonclse1.
DR InterPro; IPR004442; ExoDNase_III.
DR InterPro; IPR004808; ExoIII_xch.
DR InterPro; IPR005135; Exo_endo_phos.
DR Pfam; PF03372; Exo_endo_phos; 1.
DR TIGRFAMS; TIGR00195; exoDNase_III; 1.
DR TIGRFAMS; TIGR00633; xch; 1.
DR PROSITE; PS00726; AP_NUCLEASE_FL_1; UNKNOWN_1.
KW Exonuclease, Complete proteome.
SO SEQUENCE 256 AA; 28541 MW; 96D15554763C8B6 CRC64;

Query Match          40.0%; Score 6; DB 16; Length 256;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRAA 15
    |||||
Db 171 PESRAA 176

RESULT 30
OY2TU9 PRELIMINARY; PRT; 257 AA.
ID OY2TU9;
AC OY2TU9;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Probable exodeoxyribonuclease III protein (EC 3.1.11.2).
GN XTR44 OR RB1396 OR SBE20689.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Plasmid pSymB (megaplasmid 2).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
CX NCBI_TaxID=382;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396508; PubMed=11481431;
RA Flinn T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
RA Vorholter F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
RA Golding B., Puehler A.;
RT "The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-
RT fixing endosymbiont Sinorhizobium meliloti.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
DR EMBL; AL603646; CAC49796.1; -
DR InterPro; IPR000097; Apendonclse1.
DR InterPro; IPR004442; ExoDNase_III.
DR InterPro; IPR004808; ExoIII_xch.
DR InterPro; IPR005135; Exo_endo_phos.
DR Pfam; PF03372; Exo_endo_phos; 1.
DR TIGRFAMS; TIGR00195; exoDNase_III; 1.
DR TIGRFAMS; TIGR00633; xch; 1.
DR PROSITE; PS00726; AP_NUCLEASE_FL_1; UNKNOWN_1.
KW Hydrolyase, Plasmid, Complete proteome.
SO SEQUENCE 257 AA; 28866 MW; 162ABFBF88DC0336 CRC64;

Query Match          40.0%; Score 6; DB 16; Length 257;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRAA 15
    |||||
Db 171 PESRAA 176

RESULT 31

```

GN XII.
OS Human T-cell leukemia virus type II (HTLV-II).
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11909;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9529716; PubMed=7539968;
RA Cimiale V., D'Agostino D.M., Zotti L., Franchini G., Felber B.K.,
Chico-Bianchi L.;
RT "Expression and characterization of proteins produced by mRNAs spliced
into the x region of the human T-cell leukemia/lymphotropic virus type
II";
RL Virology 209:445-456(1995).
DR EMBL: I41677; AAA98640.1; -
SQ SEQUENCE 216 AA; 23977 MW; 44E8CE67FF85EC84 CRC64;

QY 2 HLGPFR 7
DB 20 HLGPFR 25
|||||
Query Match 40.0%; Score 6; DB 15; Length 216;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 24
Q9PX29 PRELIMINARY; PRT; 240 AA.
AC Q9PX29;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE P30 II protein.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93124536; PubMed=8419636;
RA Gessain A., Boeri E., Yanagihara R., Gallo R.C., Franchini G.;
RT "Complete nucleotide sequence of a highly divergent human T-cell
leukemia (lymphotropic) virus type I (HTLV-I) variant from melanesia:
genetic and phylogenetic relationship to HTLV-I strains from other
geographical regions.";
RT J. Virol. 67:1015-1023(1993).
SQ SEQUENCE 240 AA; 26561 MW; A17BD376692D2E37 CRC64;

QY 2 HLGPFR 7
DB 211 HLGPFR 216
|||||
Query Match 40.0%; Score 6; DB 15; Length 240;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 25
Q96I83 PRELIMINARY; PRT; 241 AA.
AC Q96I83;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE Similar to NGF-A binding protein 2 (ERG1 binding protein
2).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MUSCLE;
RA Strausberg R.;

RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC007756; AAH07756.1; -
SQ SEQUENCE 241 AA; 25927 MW; DEF799BD103ED93C CRC64;

QY 7 RSTPES 12
DB 188 RSTPES 193
|||||
Query Match 40.0%; Score 6; DB 4; Length 241;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 26
Q9PXV7 PRELIMINARY; PRT; 241 AA.
AC Q9PXV7;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE PX-TAX-ORF II fusion protein.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92409607; PubMed=1528897;
RA Koranik I.J., Gessain A., Klotman M.E., Lo Monica A., Berneman Z.N.,
Franchini G.;
RT "Protein isoforms encoded by the PX region of human T-cell
leukemia/lymphotropic virus type I.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:8813-8817(1992).
SQ SEQUENCE 241 AA; 26791 MW; 7607F1F6655CF5E CRC64;

QY 2 HLGPFR 7
DB 212 HLGPFR 217
|||||
Query Match 40.0%; Score 6; DB 15; Length 241;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 27
Q65226 PRELIMINARY; PRT; 243 AA.
AC Q65226;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE ORF 14L.
OS African swine fever virus (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
OX NCBI_TaxID=10497;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MALAWI LIL20 /1;
RX MEDLINE=94014996; PubMed=8409937;
RA Vydellingum S., Baylis S.A., Bristow C., Smith G.L., Dixon L.K.;
RT "Duplicate genes within the variable right end of the genome of a
pathogenic isolate of African swine fever virus.";
RL J. Gen. Virol. 74:2125-2130(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MALAWI LIL20 /1;
RA Dixon L.K., Twigg S.R.F., Baylis S.A., Vydellingum S., Bristow C.,
Hammond J.M., Smith G.L.;
RT "Nucleotide sequence of a 55 kbp region from the right end of the
genome of a pathogenic African swine fever virus isolate (Malawi
LIL20/1).";
RL J. Gen. Virol. 7:1655-1684(1994).
DR EMBL: X71982; CAA50817.1; -
SQ SEQUENCE 243 AA; 27569 MW; DEF79E4705EE9FE1 CRC64;


```

Query Match          40.0%; Score 6; DB 5; Length 139;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRA 14
Db 19 TPESRA 24

RESULT 20
O9AHP1 PRELIMINARY; PRT; 160 AA.
AC O9AHP1;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Lrp-regulator-like protein.
OS Agrobacterium tumefaciens.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=358;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D3;
RX MEDLINE=21322708; PubMed=11429459;
RA Trotz S., Bauer R., Knackmuss H.J., Stolz A.;
RT "Genetic and biochemical characterization of an epantiosselective
RT amidease from Agrobacterium tumefaciens strain d3."
RL Microbiology 147:1815-1824(2001).
CC -1 SIMILARITY: BELONGS TO THE ASNC FAMILY OF TRANSCRIPTOMAL
CC REGULATORS.
DR EMBL; AF315580; AAK28495.1; -.
DR InterPro; IPR000485; ASNC_trans_reg.
DR Pfam; PF01037; ASNC_trans_reg; 1.
DR PRINTS; PR00033; HTASNC.
DR SMART; SM00344; HTH_ASNC; 1.
KW DNA-binding; Transcription regulation.
SQ SEQUENCE 160 AA; 18053 MW; 3C4385D170434E49 CRC64;

Query Match          40.0%; Score 6; DB 2; Length 160;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPES 12
Db 83 RSTPES 88

RESULT 21
O9RRC1 PRELIMINARY; PRT; 168 AA.
AC O9RRC1;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE MolYbdenum cofactor biosynthesis protein C.
GN DR2571.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OC Deinococcales; Deinococcus.
OX NCBI_TaxID=129;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.S., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Matarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus

```

```

RT radiodurans R1."
RL Science 286:1571-1577(1999).
DR EMBL; AE002086; AAF12111.1; -.
DR HSSP; P30747; IEKR.
DR TIGR; DR2571; -.
DR InterPro; IPR002820; MoAC.
DR Pfam; PF01967; Moac; 1.
DR TIGRFAMs; TIGR00581; moac; 1.
KW Complete proteome.
SQ SEQUENCE 168 AA; 17433 MW; EDD6342392BBD00 CRC64;

Query Match          40.0%; Score 6; DB 16; Length 168;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRA 15
Db 44 PESRA 49

RESULT 22
O8XS36 PRELIMINARY; PRT; 186 AA.
AC O8XS36;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Probable transcription regulator protein.
GN RSP0645 OR NS03577.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM1100;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Ariat M., Billault A., Brotier P., Camus J.C., Catolico L.,
RA Chandler M., Choisne N., Claudel-Renard C., Cunac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Sautin W., Schlex T.,
RA Sigler P., Thebault P., Whalen M., Winkler P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum."
RL Nature 415:497-502(2002).
DR EMBL; AL646080; CAD17796.1; -.
DR InterPro; IPR000485; ASNC_trans_reg.
DR InterPro; IPR000345; CytC_heme_bind.
DR Pfam; PF01037; ASNC_trans_reg; 1.
DR PRINTS; PR00033; HTASNC.
DR SMART; SM00344; HTH_ASNC; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW Plasmid; Complete proteome.
SQ SEQUENCE 186 AA; 20699 MW; 603ADC806078A5F1 CRC64;

Query Match          40.0%; Score 6; DB 16; Length 186;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPES 12
Db 108 RSTPES 113

RESULT 23
O80824 PRELIMINARY; PRT; 216 AA.
AC O80824;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE Protein 28 x11.

```

RT "Complete nucleotide sequence of a highly divergent human T-cell
leukemia (lymphotropic) virus type I (HTLV-I) variant from melanesia;
RT genetic and phylogenetic relationship to HTLV-I strains from other
RT geographical regions.";
RL J. Virol. 67:1015-1023(1993).
SQ SEQUENCE 87 AA; 9988 MW; 34A239AC7EE7D9CD CRC64;

Query Match 40.0%; Score 6; DB 15; Length 87;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HLGPHR 7
|||||
DB 58 HLGPHR 63

RESULT 16
Q9PXY6 PRELIMINARY; PRT; 87 AA.
AC Q9PXY6;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Human T-lymphotropic protein.
OS Viruses; Retroviral viruses; Retroviridae; Deltaretrovirus.
OC NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92409607; PubMed=1528897;
RA Koratnik I.J., Gessain A., Klotman M.E., Lo Monico A., Berneman Z.N.,
RA Franchini G.;
RT "Protein isoforms encoded by the pX region of human T-cell
leukemia/lymphotropic virus type I";
RL Proc. Natl. Acad. Sci. U.S.A. 89:8813-8817(1992).
SQ SEQUENCE 87 AA; 10080 MW; 12DDEBDB683B277 CRC64;

Query Match 40.0%; Score 6; DB 15; Length 87;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HLGPHR 7
|||||
DB 58 HLGPHR 63

RESULT 17
Q70647 PRELIMINARY; PRT; 113 AA.
AC Q70647;
DT 01-AUG-1998 (TReMBLrel. 07, Created)
DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Hypothetical 12.9 kDa protein.
OS Simian T-lymphotropic virus 2.
OC Viruses; Retroviral viruses; Retroviridae; Deltaretrovirus.
OC NCBI_TaxID=33748;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=PP1664;
RX MEDLINE=98229455; PubMed=9568035;
RA Van Brussel M., Salemi M., Liu H.F., Gabriels J., Goubau P.,
RA Desmyter J., Vandamme A.M.;
RT "The simian T-lymphotropic virus SFLY-PP1664 from Pan paniscus is
RT distinctly related to HTLV-2 but differs in genomic organization.";
RL Virology 243:366-379(1998).
KW EMBL: Y14570; CAA74907.1; -.
DM Hypothetical protein.
FT CHAIN 5 113 POTENTIAL.
SQ SEQUENCE 113 AA; 12934 MW; 23EC3D95B410CBAE CRC64;

Query Match 40.0%; Score 6; DB 15; Length 113;
Best Local Similarity 100.0%; Pred. No. 26;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HLGPHR 7
|||||
DB 58 HLGPHR 63

RESULT 18
Q930H6 PRELIMINARY; PRT; 138 AA.
ID Q930H6
AC Q930H6;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Hypothetical protein RA0220.
GN RA0220 OR SMA0412.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OS Plasmid pSymba (megaplasmid 1).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OC NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396509; PubMed=11481432;
RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
RA Barloy-Hubler F., Bowser L., Capela D., Galibert F., Gonzy J.,
RA Gurjel M., Hong A., Huzar L., Hyman R.W., Kahn M.L.,
RA Kaiman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;
RT "Nucleotide sequence and predicted functions of the entire
RT Sinorhizobium meliloti pSymba megaplasmid.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
DR EMBL: AE007215; AAK64878.1; -.
KW Hypothetical protein; Plasmid; Complete proteome.
SQ SEQUENCE 138 AA; 15128 MW; 3E902054418D57A9 CRC64;

Query Match 40.0%; Score 6; DB 16; Length 138;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHRS 8
|||||
DB 109 LGPHRS 114

RESULT 19
Q9XX29 PRELIMINARY; PRT; 139 AA.
ID Q9XX29
AC Q9XX29;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Y39A1A.18 protein.
GN Y39A1A.18.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OC NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Wall M.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL: AL031633; CAA21011.1; -.
SQ SEQUENCE 139 AA; 15720 MW; 20782E03FC78F7F5 CRC64;

DR SMART: SM00255; TIR; 1.
KM Hypothetical protein.
SQ SEQUENCE 959 AA; 106032 MW; CEFCADEE507B76E2 CRC64;

Query Match
Best Local Similarity 46.7%; Score 7; DB 16; Length 959;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPESR 13
|||||||
DB 274 RSTPESR 280

RESULT 12

O9CRO5 PRELIMINARY; PRT; 58 AA.

DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE 16 days embryo lung cDNA, RIKEN full-length enriched library,
clone:8430430021, full insert sequence (Fragment).
GN CHD1.

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;

RP [1]

SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYONIC LUNG;

RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shitagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Aochi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Mashio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein W.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustinckin S., Hall D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Norone P., Rong B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
RA Wyszewski B., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;

RT *Functional annotation of a full-length mouse cDNA collection.*;
RL Nature 409:685-690(2001).
DR EMBL: AK018451; BAB31219.1; -.

DR MGD: MGI:88393; Chd1.

FT NON_TER 1
SQ SEQUENCE 58 AA; 6773 MW; 86F678EA2FEE67A4 CRC64;

Query Match
Best Local Similarity 40.0%; Score 6; DB 11; Length 58;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 HRSTPE 11
|||||||
DB 45 HRSTPE 50

RESULT 13

O12385 PRELIMINARY; PRT; 65 AA.

AC O12385;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE , Rex p13/p30 protein (Fragment).

GN REX.
OS Simian T-lymphotropic virus 1.
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxId=33747;

RP SEQUENCE FROM N.A.

RC STRAIN=KA;

RX MEDLINE=96120820; PubMed=9460922;

RA Verschoor E.J., Warren K., Niphuis H., Heriyanto, Swan R.A.,
RA Heeney J.L.;

RT "Characterization of a simian T-lymphotropic virus from a wild-caught
orangutan (Pongo pygmaeus) from Kalimantan, Indonesia.";
RL J. Gen. Virol. 79:51-55(1998).

DR EMBL: Y13146; CAA73607.1; -.

FT NON_TER 1
SQ SEQUENCE 65 AA; 7434 MW; 044211FEBF74AA66 CRC64;

Query Match
Best Local Similarity 40.0%; Score 6; DB 15; Length 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HLGPHR 7
|||||||
DB 36 HLGPHR 41

RESULT 14

O80794 PRELIMINARY; PRT; 87 AA.

AC O80794;
DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE X protein.

GN XII.

OS Human T-lymphotropic virus 1.

OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.

OX NCBI_TaxId=11908;

RP [1]

SEQUENCE FROM N.A.
RX MEDLINE=86232270; PubMed=2897612;

RA Tsujimoto A., Teruuchi T., Imamura J., Shimotohno K., Miyoshi I.,
RA Miwa M.;

RT "Nucleotide sequence analysis of a provirus derived from HTLV-1-
associated myelopathy (HAM)."

RL MOL. BIOL. Med. 5:29-42(1988).

DR EMBL: M37301; AAA43391.1; -.

SQ SEQUENCE 87 AA; 9963 MW; 141F3B7BA838B273 CRC64;

Query Match
Best Local Similarity 40.0%; Score 6; DB 15; Length 87;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HLGPHR 7
|||||||
DB 58 HLGPHR 63

RESULT 15

O9PX28 PRELIMINARY; PRT; 87 AA.

AC O9PX28;
DT 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE P13 II protein.

OS Human T-lymphotropic virus 1.

OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.

OX NCBI_TaxId=11908;

RP [1]

SEQUENCE FROM N.A.
RX MEDLINE=93124536; PubMed=8419636;

RA Gessain A., Boeri E., Yanagihara R., Gallo R.C., Franchini G.;

RA Matanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.,
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 Mesorhizobium loti.";
 RL DNA Res. 7:331-338(2000).
 DR EMBL: AP003007; BAB52218.1; -;
 DR InterPro: IPR000843; HTH_LactI.
 DR Pfam: PF00356; LactI.1.
 DR PRINTS: PR00036; HTHLACTI.
 DR SMART: SM00354; HTH_LACTI.1.
 DR PROSITE: PS00356; HTH_LACTI_FAMILY; UNKNOWN_1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 343 AA; 37647 MW; CF301E16D225298 CRC64;

Query Match 46.7%; Score 7; DB 16; Length 343;
 Best Local Similarity 100.0%; Pred. No. 5.4;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRA 15
 Db 285 TPESRA 291

RESULT 9
 092VZ0 PRELIMINARY; PRT; 343 AA.

ACI 092VZ0;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Putative transcriptional regulator, LactI family protein.
 GN R0557 OR SMB20817.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OG Plasmid pSymB (megaplasmid 2).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID:382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-1021;
 RX MEDLINE-21396508; PubMed-11481431;
 RA Finan T.M., Weidner S., Wong K., Buhmester J., Chain P.,
 RA Vorheeler F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
 RA Golding B., Puchler A.;
 RT "The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-
 fixing endosymbiont Sinorhizobium meliloti.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
 DR EMBL: AL603644; CAC48957.1; -;
 DR InterPro: IPR000843; HTH_LactI.
 DR Pfam: PF00356; LactI.1.
 DR PROSITE: PS00356; HTH_LACTI_FAMILY; UNKNOWN_1.
 KW Plasmid; Hypothetical protein; Complete proteome.
 SQ SEQUENCE 343 AA; 37381 MW; 714FF82A41E4522D CRC64;

Query Match 46.7%; Score 7; DB 16; Length 343;
 Best Local Similarity 100.0%; Pred. No. 5.4;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRA 15
 Db 285 TPESRA 291

RESULT 10
 08S6U7 PRELIMINARY; PRT; 738 AA.

ACI 08S6U7;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Putative retrotransposable elements TNP2.

GN OSJNB0014J14.22.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukarya; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID:39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wing R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,
 RA Saski C., Henry D., Oates R., Simmons J.;
 RT "rice Genomic Sequence.";
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC092172; AAM18162.1; -;
 SQ SEQUENCE 738 AA; 83826 MW; 0626A5AC35C4243B CRC64;

Query Match 46.7%; Score 7; DB 10; Length 738;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPESRA 14
 Db 146 STPESRA 152

RESULT 11
 09RL12 PRELIMINARY; PRT; 959 AA.

ACI 09RL12;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical protein SC00305.
 GN SC00305 OR SC569.14.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
 OX NCBI_TaxID:1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Seeger K.J., Harris D.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA James K.D., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RX MEDLINE-97000351; PubMed-8843436;
 RA Redenbach M., Kleser H.M., Denapate D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2) / M145;
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Crozin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 coelicolor A3(2)." (2002).
 RL Nature 417:141-147(2002).
 DR EMBL: AL11385; CAB55662.1; -;
 DR InterPro: IPR000157; TIR_domain.

OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
 OX NCBI_TaxID=1502;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Brown S.P., Murphy L.D., Harris D.;
 RL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA MEDLINE=98241550; PubMed=9573173;
 RX Redenbach M., Ikeda K., Yamasaki M., Kinashi H.;
 RT "Cloning and physical mapping of the EcoRI fragments of the giant
 RT linear plasmid SCP1.";
 RL J. Bacteriol. 180:2796-2799(1998).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowicz E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 RT coelicolor A3(2).";
 RL Nature 417:141-147(2002).
 DR EMBL: AL590464; CAC36767.1; -;
 DR InterPro: IPR001031; Thioesterase.
 DR Pfam: PF00975; Thioesterase; 1.
 DR Plasmid.
 SQ SEQUENCE 269 AA; 28898 MW; 2FAEDA79BEA0867 CRC64;
 Query Match 46.7%; Score 7; DB 16; Length 269;
 Best Local Similarity 100.0%; Pred. No. 4.4;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 9 TPESRAA 15
 Db 234 TPESRAA 240
 RESULT 6
 OBR2A2 PRELIMINARY; PRT; 316 AA.
 AC OBR2A2;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Vomeronasal receptor VIRE8.
 GN VIRE8.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129X1/SVJ;
 RA MEDLINE=21676859; PubMed=11802169;
 RA Rodriguez I., Del Punta K., Rothman A., Ishii T., Mombaerts P.;
 RT "Multiple new and isolated families within the mouse superfamily of
 RT V1R vomeronasal receptors";
 RL Nat. Neurosci. 5:134-140(2002).
 DR EMBL: AY065508; AAL47913.1; -;
 KW * Receptor.

SQ SEQUENCE 316 AA; 36450 MW; 3C08CBDA5992C5F CRC64;
 Query Match 46.7%; Score 7; DB 11; Length 316;
 Best Local Similarity 100.0%; Pred. No. 5.1;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 8 TPESRAA 14
 Db 243 TPESRAA 249
 RESULT 7
 O9A5T8 PRELIMINARY; PRT; 327 AA.
 ID O9A5T8;
 AC O9A5T8;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-OCT-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE N4-(beta-N-acetylglucosaminyl)-L-asparaginase, putative.
 GN CC2359.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
 OC Caulobacter.
 OX NCBI_TaxID=155892;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 19089 / CB15;
 RX MEDLINE=21173698; PubMed=11259647;
 RA Nieman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 RA Pollock A., Nelson W.C., Newton A., Stephens C., Padke N.D., Ely B.,
 RA Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
 RA Uitterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of Caulobacter crescentus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 DR EMBL: AE005905; AK24330.1; -;
 DR HSP: Q47898; IAVY.
 DR MEROPS: T02.001; -;
 DR TIGR: CC2359; -;
 DR InterPro: IPR000246; Asparaginase_2.
 DR Pfam: PF01112; Asparaginase_2; 1.
 KW Complete proteome.
 SQ SEQUENCE 327 AA; 34165 MW; 66369845ACDD414 CRC64;
 Query Match 46.7%; Score 7; DB 16; Length 327;
 Best Local Similarity 100.0%; Pred. No. 5.2;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 9 TPESRAA 15
 Db 139 TPESRAA 145
 RESULT 8
 O9BAV4 PRELIMINARY; PRT; 343 AA.
 ID O9BAV4;
 AC O9BAV4;
 DT 01-OCT-2001 (TREMBLrel. 18, Created)
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical protein ml15838.
 GN ML15838.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAF303099;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,

DR PRINTS; PRO1217; PRICHEXTENSIN.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 347 AA; 34908 MW; 1C3CEA9C1EAE4C CRC64;

Query Match 53.3%; Score 8; DB 16; Length 347;
Best Local Similarity 100.0%; Pred. No. 0.45;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPESRA 14
DB 274 RSTPESRA 281

RESULT 2
ID 090X08 PRELIMINARY; PRT; 468 AA.

AC 090X08; 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, last sequence update)
DE 01-MAR-2002 (TREMBlrel. 20, last annotation update)
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL4; TISSUE=BRAIN;
RA Korade Mirnics Z., Keryanov S., Lovelock J., Corey S.J.;
RT "Cloning of chicken presentinins."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY043492; AAK95408.1; -;
DR InterPro: IPR001108; Presentinlin.
DR Pfam: PF01080; Presentinlin.
DR PRINTS; PRO1072; PRESENTINLIN.
SQ SEQUENCE 468 AA; 52812 MW; B746BDA3BAC0BFA CRC64;

Query Match 53.3%; Score 8; DB 13; Length 468;
Best Local Similarity 100.0%; Pred. No. 0.57;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPESRAA 15
DB 354 STPESRAA 361

RESULT 3
ID 0905H6 PRELIMINARY; PRT; 19 AA.

AC 0905H6; 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, last annotation update)
OS Human immunodeficiency virus type 1.
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97CG152;
RA Tanguich Y., Takehsia J., Bikandou B., Mboudjeka I.,
RA N'Doundou N'kodia M.-Y., M'pand M., M'Pele P., Harada Y., Hayami M.,
RA Ichimura H., Paria H.-J.;
RT "Genetic Subtypes of HIV Type 1 Based on the vpu/env Sequences in
RT Republic of Congo-Brazzaville."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AA410445; AAL10246.1; -;
FT NON_TER 1 19
FT NON_TER 1 19
SQ SEQUENCE 19 AA; 2306 MW; 366046FD0069372 CRC64;

Query Match 46.7%; Score 7; DB 15; Length 19;

Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPESR 13
DB 4 RSTPESR 10

RESULT 4
ID 09JN88 PRELIMINARY; PRT; 241 AA.

AC 09JN88; 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)
DE 01-OCT-2001 (TREMBlrel. 18, last annotation update)
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Bruton C.J., Wietzorrek A., Hartley N., Woodburn L., Chater K.F.;
RT "Genes involved in methylenomycin biosynthesis from plasmid SCP1 of
RT Streptomyces coelicolor A3(2)."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=88112873; PubMed=2828187;
RA Neal R.J., Chater K.F.;
RT "Nucleotide sequence analysis reveals similarities between proteins
RT determining methylenomycin A resistance in Streptomyces and
RT tetracycline resistance in eubacteria."
RL Gene 58:229-241(1987).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=85284984; PubMed=2992952;
RA Chater K.F., Bruton C.J.;
RT "Resistance, regulatory and production genes for the antibiotic
RT methylenomycin are clustered."
RL EMBL J. 4:1893-1897(1985).
DR EMBL: AJ276673; CAB82867.1; -;
DR InterPro: IPR001031; Thioesterase.
DR Pfam: PF00975; Thioesterase; 1.
SQ SEQUENCE 241 AA; 25801 MW; 55D39777F472DDBA CRC64;

Query Match 46.7%; Score 7; DB 2; Length 241;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRAA 15
DB 206 TPESRAA 212

RESULT 5
ID 09ACS2 PRELIMINARY; PRT; 269 AA.

AC 09ACS2; 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, last annotation update)
OS Putative thioesterase, Mmyt.
GN MMYT OR SCP1.241C.
OS Streptomyces coelicolor.
OG Plasmid SCP1.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

90	5	33.3	56	16	Q9W280	Q9w280 thermotoga
91	5	33.3	62	4	Q14274	Q14274 homo sapien
92	5	33.3	62	12	Q14274	Q14274 chilo iride
93	5	33.3	69	16	Q8XB2	Q8xb2 escherichia
94	5	33.3	71	16	Q8XK0	Q8xk0 ralslonia s
95	5	33.3	72	12	Q9E796	Q9e796 hepaticis c
96	5	33.3	76	17	Q9YB07	Q9yeb07 aeropyrum p
97	5	33.3	78	7	Q8WM11	Q8wm11 rangifer ta
98	5	33.3	78	7	Q8WM08	Q8wm08 rangifer ta
99	5	33.3	83	7	Q19220	Q19220 rangifer ta
100	5	33.3	83	7	Q19221	Q19221 rangifer ta
101	5	33.3	83	7	Q19222	Q19222 rangifer ta
102	5	33.3	83	7	Q19227	Q19227 rangifer ta
103	5	33.3	85	6	Q95MA6	Q95ma6 equus caball
104	5	33.3	87	16	Q96AZ6	Q96az6 homo sapien
105	5	33.3	87	16	Q8YFP6	Q8yfp6 brucella me
106	5	33.3	91	6	Q9N001	Q9n001 macaca fasc
107	5	33.3	92	2	Q50101	Q50101 mycobacteri
108	5	33.3	96	5	Q9N8N8	Q9n8n8 trypanosoma
109	5	33.3	96	9	Q8SC82	Q8sc82 stx2 convey
110	5	33.3	97	13	Q90WF2	Q90wf2 paralicthy
111	5	33.3	105	16	Q82DG2	Q82dg2 yerstinia pe
112	5	33.3	106	4	Q99494	Q99494 homo sapien
113	5	33.3	109	10	Q9FUT8	Q9fut8 zea mays (m
114	5	33.3	110	2	Q8VNV2	Q8vvn2 vibrio chol
115	5	33.3	110	10	Q9FPT7	Q9fpt7 zea mays (m
116	5	33.3	112	11	Q9D503	Q9d503 mus musculu
117	5	33.3	112	12	Q9YR60	Q9yr60 aleutian m
118	5	33.3	112	16	Q8XV40	Q8xv40 ralslonia s
119	5	33.3	113	10	P81170	P81170 trifolium r
120	5	33.3	117	11	Q92304	Q92304 cavia porce
121	5	33.3	118	17	Q974P5	Q974p5 sulfolobus
122	5	33.3	119	12	Q11311	Q11311 moluscum c
123	5	33.3	121	17	Q8TIX0	Q8tix0 methanosarc
124	5	33.3	127	16	Q92RR8	Q92rr8 rhizobium m
125	5	33.3	127	16	Q8YEG0	Q8yeg0 brucella me
126	5	33.3	130	15	Q88013	Q88013 chimpanzee
127	5	33.3	131	15	Q90EX7	Q90ex7 simian immu
128	5	33.3	131	15	Q90EX2	Q90ex2 simian immu
129	5	33.3	134	16	Q8X104	Q8x104 clostridium
130	5	33.3	135	2	Q52214	Q52214 pseudomonas
131	5	33.3	135	16	Q9PGJ5	Q9pgj5 xyella fas
132	5	33.3	138	11	Q9CXR5	Q9cxr5 mus musculu
133	5	33.3	139	4	Q8WY52	Q8wy52 homo sapien
134	5	33.3	139	10	Q8S071	Q8s071 oryza sativ
135	5	33.3	139	13	P79799	P79799 micrurus co
136	5	33.3	143	5	Q9NM50	Q9nm50 leishmania
137	5	33.3	143	10	Q942D1	Q942d1 oryza sativ
138	5	33.3	145	2	Q823C2	Q823c2 synecococc
139	5	33.3	145	16	Q8YR84	Q8yr84 anabena sp
140	5	33.3	148	9	Q9MCM1	Q9mcm1 streptococc
141	5	33.3	150	4	Q9Y413	Q9y413 homo sapien
142	5	33.3	150	11	Q923U3	Q923u3 cavia porce
143	5	33.3	150	17	Q8T5I3	Q8t5i3 methanosarc
144	5	33.3	152	2	Q8R0P5	Q8r0p5 thermus the
145	5	33.3	154	2	Q9RE11	Q9rell micrococcus
146	5	33.3	154	2	Q47394	Q47394 escherichia
147	5	33.3	154	12	Q9DS07	Q9ds07 foot-and-mo
148	5	33.3	155	10	Q48591	Q48591 lactococcus
149	5	33.3	155	10	Q9W9A2	Q9w9a2 arabidopsis
150	5	33.3	157	16	Q8UE78	Q8ue78 agrobacteri
151	5	33.3	159	2	Q9F7V2	Q9f7v2 staphylococ
152	5	33.3	159	2	Q8VSV5	Q8vsv5 staphylococ
153	5	33.3	160	10	Q04218	Q04218 brassica na
154	5	33.3	160	2	Q9F484	Q9f484 alteromonas
155	5	33.3	161	17	Q58905	Q58905 pyrococcus
156	5	33.3	161	6	Q9N113	Q9n113 entiemur ful
157	5	33.3	163	2	Q53000	Q53000 escherichia
158	5	33.3	163	17	Q96129	Q96129 homo sapien
159	5	33.3	163	17	Q9H508	Q9h508 halobacteri
160	5	33.3	165	10	Q24121	Q24121 nicotiana p
161	5	33.3	165	16	Q9CCZ0	Q9ccz0 mycobacteri
162	5	33.3	167	10	Q9LH78	Q9lh78 arabidopsis

163	5	33.3	168	3	Q06236	Q06236 saccharomyc
164	5	33.3	168	5	Q97194	Q97194 leishmania
165	5	33.3	170	4	Q8WYG9	Q8wvg9 homo sapien
166	5	33.3	170	4	Q96CC4	Q96cc4 homo sapien
167	5	33.3	172	2	Q52386	Q52386 lactobacill
168	5	33.3	173	16	Q9RR06	Q9rr06 delnoccocus
169	5	33.3	174	13	Q9W6F2	Q9w6f2 gallus gall
170	5	33.3	175	2	P95548	P95548 pseudomonas
171	5	33.3	176	5	Q9VKY6	Q9vky6 drosophila
172	5	33.3	176	11	Q80ZT0	Q80zt0 mus musculu
173	5	33.3	181	4	Q00441	Q00441 homo sapien
174	5	33.3	181	16	Q92LB0	Q92lb0 rhizobium m
175	5	33.3	181	16	Q9LOE7	Q9loe7 streptomyce
176	5	33.3	183	4	Q9BTA6	Q9bta6 homo sapien
177	5	33.3	183	16	Q82013	Q82013 anabena sp
178	5	33.3	185	10	Q81624	Q81624 arabidopsis
179	5	33.3	187	16	Q9KYZ4	Q9kyz4 streptomyc
180	5	33.3	188	11	Q9ERS8	Q9ers8 raltus norv
181	5	33.3	189	10	Q9ZUE1	Q9zue1 arabidopsis
182	5	33.3	189	16	Q8XAL7	Q8xal7 escherichia
183	5	33.3	190	11	P70368	P70368 mus musculu
184	5	33.3	191	11	Q9J329	Q9j329 mus musculu
185	5	33.3	191	16	Q9PXR8	Q9pxr8 xyella fas
186	5	33.3	192	8	Q95C95	Q95c95 nupertia lu
187	5	33.3	195	10	Q9L113	Q9l113 arabidopsis
188	5	33.3	196	8	Q95OR1	Q95or1 spizellomyc
189	5	33.3	197	4	Q9U0R9	Q9u0r9 homo sapien
190	5	33.3	197	17	Q9Y974	Q9y974 aeropyrum p
191	5	33.3	201	4	Q8H6I6	Q8h6i6 homo sapien
192	5	33.3	201	10	Q9C7M6	Q9c7m6 arabidopsis
193	5	33.3	201	12	Q8UY82	Q8uy82 simian aden
194	5	33.3	201	16	Q8UI48	Q8ui48 agrobacteri
195	5	33.3	203	10	Q81611	Q81611 carica papa
196	5	33.3	203	11	Q9JH10	Q9jh10 mus musculu
197	5	33.3	204	5	Q20551	Q20551 caenohabdi
198	5	33.3	204	5	Q9N709	Q9n709 leishmania
199	5	33.3	206	8	Q35872	Q35872 lycopersico
200	5	33.3	206	8	Q79373	Q79373 brassica na

ALIGNMENTS

RESULT 1					
ID	Q9RR94	PRELIMINARY;	PRT;	347	AA.
AC	Q9RR94;				
DT	01-MAY-2000 (TREMBlrel. 13, Created)				
DT	01-MAY-2000 (TREMBlrel. 13, Last sequence update)				
DE	01-MAR-2002 (TREMBlrel. 20, Last annotation update)				
OS	Hypothetical protein DR2600.				
OC	Deinococcus radiodurans.				
CC	Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;				
OX	NCBI_TaxID=1299;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-RI:				
RA	MEDLINE-20036896; PubMed-10567266;				
RA	White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,				
RA	Dodson R.D., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,				
RA	Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,				
RA	Vamathevan J.J., Lam P., McDonald L., Uterback T., Zaleski R.C.,				
RA	Makerova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,				
RA	Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,				
RA	Fraser C.M.,				
RT	"Genome sequence of the radioresistant bacterium Deinococcus				
RT	radiodurans RI.";				
RL	Science 286:1571-1577(1999).				
DR	EMBL; AE002089; AAF12144.1; -.				
DR	TIGR; DR2600; -.				
DR	InterPro; IPR002965; P_Rich_extensn.				

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2003, 14:21:50 ; Search time 69 Seconds
(without alignments)
44.793 Million cell updates/sec

Title: US-09-689-159a-2_COPY_346_360
Perfect score: 15
Sequence: 1 SHLGHRSPTESRAA 15

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size : 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 200 summaries

Database :

SPREMBL_21:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.rodent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriap:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	53.3	347	16	Q9RR94	Q9RR94 deinococcus
2	53.3	468	13	Q90X08	Q90X08 gallus gall
3	46.7	19	15	Q905H6	Q905H6 human immun
4	46.7	241	2	Q9JN88	Q9JN88 streptomyc
5	46.7	269	16	Q9ACS2	Q9ACS2 streptomyc
6	46.7	316	11	Q8R2A2	Q8R2A2 mus musculu
7	46.7	327	16	Q9A5T8	Q9A5T8 caulobacter
8	46.7	343	16	Q9A5V4	Q9A5V4 rhizobium l
9	46.7	738	10	Q92VZ0	Q92VZ0 oryza sativ
10	46.7	738	10	Q8S6U7	Q8S6U7 oryza sativ
11	46.7	959	16	Q9RLI2	Q9RLI2 streptomyc
12	40.0	58	11	Q9CRO5	Q9CRO5 mus musculu
13	40.0	65	15	Q12385	Q12385 simian t-ly
14	40.0	87	15	Q80794	Q80794 human t-lym
15	40.0	87	15	Q9PXZ8	Q9PXZ8 human t-lym
16	40.0	87	15	Q9PXV6	Q9PXV6 human t-lym

17	40.0	113	15	Q70647	Q70647 simian t-ly
18	40.0	138	16	Q930H6	Q930H6 rhizobium m
19	40.0	139	5	Q9XX29	Q9XX29 caenorhabd
20	40.0	160	2	Q9AHF1	Q9AHF1 agrobacteri
21	40.0	168	16	Q9RRC1	Q9RRC1 deinococcus
22	40.0	186	16	Q8XS36	Q8XS36 raietonia s
23	40.0	216	15	Q80824	Q80824 human t-cel
24	40.0	240	15	Q9PXZ9	Q9PXZ9 human t-lym
25	40.0	241	4	Q96183	Q96183 homo sapien
26	40.0	241	15	Q9PXV7	Q9PXV7 human t-lym
27	40.0	243	12	Q65226	Q65226 african swi
28	40.0	255	16	Q8Y0G6	Q8Y0G6 raietonia s
29	40.0	256	16	Q98J74	Q98J74 rhizobium l
30	40.0	257	16	Q92TU9	Q92TU9 rhizobium m
31	40.0	259	6	Q9BDK8	Q9BDK8 sus scrofa
32	40.0	262	16	Q9A6S2	Q9A6S2 caulobacter
33	40.0	301	16	Q9R146	Q9R146 streptomyc
34	40.0	315	10	Q86J99	Q86J99 oryza sativ
35	40.0	343	10	Q9XEP5	Q9XEP5 sorghum bic
36	40.0	365	12	Q91T06	Q91T06 lupula herp
37	40.0	389	3	Q9HEP7	Q9HEP7 blumeria gr
38	40.0	401	16	Q07569	Q07569 bacillus su
39	40.0	407	4	Q9UFE4	Q9UFE4 homo sapien
40	40.0	408	16	Q9AB45	Q9AB45 caulobacter
41	40.0	413	16	Q8Y1W1	Q8Y1W1 raietonia s
42	40.0	426	16	Q99XK9	Q99XK9 streptococc
43	40.0	429	16	Q97NC9	Q97NC9 streptococc
44	40.0	444	10	Q95R06	Q95R06 arbidopsis
45	40.0	468	5	Q95007	Q95007 caenorhabd
46	40.0	563	11	Q925F7	Q925F7 raietius norv
47	40.0	575	16	Q9RJQ2	Q9RJQ2 streptomyc
48	40.0	577	11	Q9D221	Q9D221 mus musculu
49	40.0	614	10	Q9FL63	Q9FL63 arbidopsis
50	40.0	617	2	Q68077	Q68077 rhodobacter
51	40.0	637	17	Q8TJB4	Q8TJB4 mechanosarc
52	40.0	640	16	Q8Y1Y1	Q8Y1Y1 anaerobac sp
53	40.0	678	4	Q8TES7	Q8TES7 homo sapien
54	40.0	735	13	Q8QGX3	Q8QGX3 xenopus lae
55	40.0	787	12	Q9E352	Q9E352 maize negro
56	40.0	805	12	Q98Y45	Q98Y45 turkey herp
57	40.0	808	12	Q98Y44	Q98Y44 turkey herp
58	40.0	808	10	Q94GT7	Q94GT7 oryza sativ
59	40.0	818	12	Q88485	Q88485 tomato bush
60	40.0	818	12	P89211	P89211 tomato bush
61	40.0	818	12	Q96606	Q96606 artichoke m
62	40.0	818	12	Q9MJD6	Q9MJD6 cymbidium r
63	40.0	851	12	Q66101	Q66101 carnation l
64	40.0	858	4	Q9N053	Q9N053 homo sapien
65	40.0	865	12	Q83291	Q83291 marek disea
66	40.0	888	16	Q9RD32	Q9RD32 streptomyc
67	40.0	889	12	Q98280	Q98280 molluscum c
68	40.0	938	10	Q8SFP5	Q8SFP5 oryza sativ
69	40.0	970	16	Q98BW9	Q98BW9 rhizobium l
70	40.0	1048	13	P79749	P79749 fuigu rubrip
71	40.0	1097	11	Q8R406	Q8R406 raietius norv
72	40.0	1117	10	Q9LRY7	Q9LRY7 arbidopsis
73	40.0	1680	10	Q64588	Q64588 arbidopsis
74	40.0	2139	5	Q9VUQ9	Q9VUQ9 drosophila
75	40.0	2596	4	Q96RNT	Q96RNT homo sapien
76	40.0	2664	5	Q26033	Q26033 plasmodium
77	40.0	2696	4	Q96L73	Q96L73 homo sapien
78	40.0	4340	2	Q30764	Q30764 streptomyc
79	33.3	23	11	P70635	P70635 raietius norv
80	33.3	26	10	Q22463	Q22463 hordeum vul
81	33.3	39	4	Q14559	Q14559 homo sapien
82	33.3	39	15	Q55267	Q55267 human t-cel
83	33.3	39	15	Q55268	Q55268 human t-cel
84	33.3	39	15	Q55269	Q55269 human t-cel
85	33.3	39	15	Q55270	Q55270 human t-cel
86	33.3	39	15	Q82210	Q82210 human t-lym
87	33.3	39	15	Q82239	Q82239 human t-cel
88	33.3	51	16	Q929G4	Q929G4 chlamydia p
89	33.3	55	2	Q939M5	Q939M5 aeromonas s

THIS PAGE BLANK (USPTO)


```

GN RECB OR CT639.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DM-3/CK;
RX MEDLINE=9900809; PubMed=9784136;
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
  Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
  Davis R.W.;
RA "Genome sequence of an obligate intracellular pathogen of humans:
  Chlamydia trachomatis."
RL Science 282:754-759(1998).
CC -1- FUNCTION: REQUIRED FOR EFFICIENT DNA REPAIR; IT CATALYZES THE
  UNWINDING OF DOUBLE-STRANDED DNA AND THE CLEAVAGE OF SINGLE-
  STRANDED DNA AND IT STIMULATES LOCAL GENETIC RECOMBINATION.
CC ALL OF THESE ACTIVITIES REQUIRE CONCOMITANT HYDROLYSIS OF ATP
  (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: Exonucleolytic cleavage (in the presence of
  ATP) in either 5'- to 3'- or 3'- to 5'- direction to yield 5'-
  phosphooligonucleotides.
CC -1- SUBUNIT: CONSIST OF THREE SUBUNITS; RECB, RECC AND RECD
  (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE HELICASE FAMILY. UVRD SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001334; AAC68243.1; -.
DR InterPro: IPR004586; RECB.
DR InterPro: IPR000212; UVRD-helicase.
DR Pfam: PF00580; UVRD-helicase; 1.
DR TIGRFAMs: TIGR00609; recb; 1.
KW Hydrolyase; Nuclease; Exonuclease; Endonuclease; Helicase; ATP-binding;
  DNA repair; Complete proteome.
FT NP_BIND 21 28 ATP (POTENTIAL).
FT SEQUENCE 1026 AA; 117828 MW; 63000D548C0F7A33 CRC64;
SQ
Query Match 33.3%; Score 5; DB 1; Length 1026;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
DB 436 RSTPE 440

RESULT 199
YG35_YEAST STANDARD: PRT; 1036 AA.
ID YG35_YEAST
AC P53273;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical 117.0 kDa protein in ASN2-PHB1 intergenic region.
GN YG8125W OR G6362.
OS Saccharomycetes; Saccharomycetaceae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
  Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / FY1679;
RX MEDLINE=97197982; PubMed=9046098;
RA "An 18.3 kb DNA fragment from yeast chromosome VII carries four
  unknown open reading frames, the gene for an Asn synthase, remnants

```

```

RT of Ty and three tRNA genes."
RL Yeast 13:171-176(1997).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: CONTAINS 1 STAS DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X83099; CAAS8161.1; -.
DR EMBL: Z72910; CAAG7136.1; -.
DR SGD: S0003357; YGR125W.
DR InterPro: IPR002645; STAS.
DR InterPro: IPR001902; Sulfate_transp.
DR Pfam: PF00916; Sulfate_transp; 1.
DR PROSITE: PS50801; STAS; 1.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 214 234
FT TRANSMEM 237 257 POTENTIAL.
FT TRANSMEM 270 290 POTENTIAL.
FT TRANSMEM 301 321 POTENTIAL.
FT TRANSMEM 339 359 POTENTIAL.
FT TRANSMEM 376 396 POTENTIAL.
FT TRANSMEM 406 426 POTENTIAL.
FT TRANSMEM 474 494 POTENTIAL.
FT TRANSMEM 516 536 POTENTIAL.
FT TRANSMEM 547 567 POTENTIAL.
FT TRANSMEM 569 589 POTENTIAL.
FT TRANSMEM 605 625 POTENTIAL.
FT TRANSMEM 665 685 POTENTIAL.
FT DOMAIN 660 781 STAS.
FT SEQUENCE 1036 AA; 116970 MW; 91A50225CA7FECB1 CRC64;
SQ
Query Match 33.3%; Score 5; DB 1; Length 1036;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGP 5
DB 240 SHLGP 244

RESULT 200
TTA4_HUMAN STANDARD: PRT; 1038 AA.
ID TTA4_HUMAN
AC P13612;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Integrin alpha-4 precursor (Integrin alpha-IV) (VLA-4) (CD49d).
GN ITGA4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89356603; PubMed=2788572;
RA Takada Y., Ellices M.J., Crouse C., Hemler M.E.;
RT "The primary structure of the alpha 4 subunit of VLA-4: homology to
  other integrins and a possible cell-cell adhesion function."
RL EMBO J. 8:1361-1368(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92201303; PubMed=1551405;
RA Rubio M., Nueda A., Vaya A., Corbi Lopez A.L.;
RT "A single mRNA encodes the alpha 150 and alpha 80/70 forms of the
  alpha subunit of VLA4."

```

RC STRAIN-Berkeley;
 RX MEDLINE-20196006; PubMed-10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sultón G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Bailev R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley S.E.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokva D., Botchan M.R., Bouck J., Brockstein P., Brotler P.,
 RA Burris K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,
 RA Cherry J.M., Chew J.S., Dahlke C., Davenport L.B., Davies P.,
 de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glöckner A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Iregyan C.,
 RA Jialili M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclib J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveril J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE OF 626-944 FROM N.A.
 RC TISSUE-Ovary;
 RX MEDLINE-99168769; PubMed-10071211;
 RA Caggese C., Ragone G., Perini B., Moschetti R., de Pinto V.,
 RA Calzari R., Baraniti P.;
 RT "Identification of nuclear genes encoding mitochondrial proteins:
 RT isolation of a collection of D. melanogaster cDNAs homologous to
 RT sequences in the Human Gene Index database.";
 RL Mol. Gen. Genet. 261:64-70(1999).
 CC -1- SUBCELLULAR LOCATION: Mitochondrial.
 CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
 CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO
 CC FRAMESHIFTS IN POSITIONS 920 AND 930.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AEO03464; AAF47217.1;
 DR EMBL: Y10912; CAA71853.1; ALT_FRAME.
 DR FlyBase: FBgn0019886; CG4598.
 DR InterPro: IPR002048; EF-hand.
 DR Pfam: PF00036; efhand; 2.
 DR SMART: SM00054; Efh; 2.
 DR PROSITE: PS00018; EF_HAND; 2.
 KW Mitochondrion; Calcium-binding; Repeat.
 FT CA_BIND 700 711 EF_HAND 1 (POTENTIAL).
 FT CA_BIND 783 794 EF_HAND 2 (POTENTIAL).
 FT CA_BIND

FT CONFLICT 687 687 E -> Q (IN REF. 2).
 FT CONFLICT 690 690 K -> G (IN REF. 2).
 FT CONFLICT 717 717 Q -> H (IN REF. 2).
 FT CONFLICT 736 736 E -> D (IN REF. 2).
 FT CONFLICT 740 740 K -> R (IN REF. 2).
 FT CONFLICT 761 761 K -> E (IN REF. 2).
 FT CONFLICT 943 943 G -> S (IN REF. 2).
 SQ SEQUENCE 1013 AA; 113579 MW; 8C710380263F262E CRC64;
 Query Match 33.3%; Score 5; DB 1; Length 1013;
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 11 ESRA 15
 Db 615 ESRA 619
 RESULT 197
 MANA_RHOMR STANDARD; PRT; 1021 AA.
 ID MANA_RHOMR
 AC P49425;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Mannan endo-1,4-beta-mannosidase (EC 3.2.1.78).
 GN MANA.
 OS Rhodothermus marinus (Rhodothermus obamensis).
 OC Bacteria; Bacteroidetes; Sphingobacteriales;
 OC Crenitrichaceae; Rhodothermus.
 ON NCBI_TaxID=29549;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 43812;
 RA Pollitz O., Krah M., Borriess R.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-beta-D-mannosidic
 CC linkages in mannans, galactomannans, glucomannans, and
 CC galactoglucomannans.
 CC -1- SIMILARITY: BELONGS TO FAMILY 26 OF GLYCOSYL HYDROLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X90947; CAA62442.1;
 DR InterPro: IPR000805; Glyco_hydro_26.
 DR Pfam: PF02156; Glyco_hydro_26; 1.
 DR PRINTS: PR00739; GLYH2RLASE26.
 KW Hydrolase; Glycosidase.
 SQ SEQUENCE 1021 AA; 115791 MW; A8F7B63109E0F715 CRC64;
 Query Match 33.3%; Score 5; DB 1; Length 1021;
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 GPHRS 8
 Db 3 GPHRS 7
 RESULT 198
 EX5B_CHLTR STANDARD; PRT; 1026 AA.
 ID EX5B_CHLTR
 AC 084645;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Exodeoxyribonuclease V beta chain (EC 3.1.11.5).
 DE

```

FT CONFLICT 121 135 PSGEKOVPPILDKQ -> AQREAGASHPPWOT (IN
FT CONFLICT 204 204 REF. 1).
FT CONFLICT 385 411 A -> G (IN REF. 1).
OHIDRSDJNRGCPPTIVEMILPWVL -> STST (IN
REF. 1).
FT CONFLICT 681 681 K -> M (IN REF. 1).
FT CONFLICT 705 705 Q -> H (IN REF. 1).
FT CONFLICT 728 728 E -> D (IN REF. 1).
FT CONFLICT 807 807 A -> V (IN REF. 1).
FT CONFLICT 873 873 O -> R (IN REF. 1).
FT CONFLICT 888 888 K -> O (IN REF. 1).
FT CONFLICT 922 977 VKRVCSEKSEKVVEDVPIPIREKHAQEDSSIDYDLS
PDTIVAHEDVYVTHL -> GQESVLLQVEGGGWMEDNRP
YTKGRNTPRCEDESKHKIMI (IN REF. 1).
SQ SEQUENCE 977 AA; 111847 MW; 6F86DA95261E0ECD CRC64;

Query Match 33.3%; Score 5; DB 1; Length 977;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGP 5
DB 500 SHLGP 504

RESULT 195
TRP4_BOVIN STANDARD; PRT; 981 AA.
ID TRP4_BOVIN STANDARD; PRT; 981 AA.
AC P79100;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Short transient receptor potential channel 4 (Trp4) (Capacitative
calcium entry channel 1) (CCE1).
GN TRPC4.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM ALPHA).
RC TISSUE-Retina, and Adrenal gland;
RX MEDLINE=97102798; PubMed=8947038;
RA Philipp S., Cavalle A., Freichel M., Wissenbach U., Zimmer S.,
RA Trost C., Marcquart A., Murakami M., Flockerzi V.;
RT "A mammalian capacitative calcium entry channel homologous to
RT Drosophila TRP and TRPL.";
RL EMBO J. 15:6166-6171(1996).
RN [2]
RN SEQUENCE FROM N.A. (ISOFORM BETA).
RC TISSUE-Adrenal gland;
RX MEDLINE=98158580; PubMed=9498815;
RA Freichel M., Wissenbach U., Philipp S., Flockerzi V.;
RT "Alternative splicing and tissue specific expression of the 5'
RT truncated bCCE 1 variant bCCE Idelta54.";
RL FEBS Lett. 422:354-358(1998).
CC -1- FUNCTION: THOUGHT TO FORM A RECEPTOR-ACTIVATED NON-SELECTIVE
CC CALCIUM PERMEANT CATION CHANNEL. PROBABLY IS OPERATED BY A
CC PHOSPHATIDYLINOSITOL SECOND MESSENGER SYSTEM ACTIVATED BY RECEPTOR
CC THIOSINE KINASES OR G-PROTEIN COUPLED RECEPTORS. HAS ALSO BEEN
CC SHOWN TO BE CALCIUM-SELECTIVE. MAY ALSO BE ACTIVATED BY
CC INTRACELLULAR CALCIUM STORE DEPLETION.
CC -1- SUBUNIT: ISOFORM ALPHA ASSOCIATES WITH INOSITOL-1,4,5-TRIPHOSPHATE
CC RECEPTOR (IP3R). INTERACTS WITH NHERF (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: ALPHA (SHOWN HERE) AND
CC BETA/DELTA 514; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ADRENAL GLAND. LOWER EXPRESSION
CC IN HEART AND RETINA. ALSO EXPRESSED IN TESTIS. THE SHORT ISOFORM
CC IS SPECIFICALLY EXPRESSED IN THE ADRENAL GLAND.
CC -1- SIMILARITY: BELONGS TO THE TRANSIENT RECEPTOR FAMILY. STRPC
CC SUBFAMILY.

```

```

CC -1- SIMILARITY: CONTAINS 2 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X99792; CA68125.1; -.
DR EMBL; AJ224862; CA12161.1; -.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR002111; Cat_channel_TrpL.
DR InterPro; IPR000636; M-channel_nlg.
DR InterPro; IPR002153; Trans_recep.
DR InterPro; IPR004729; Trp_Catchannel.
DR Pfam; PF00023; ank; 2.
DR Pfam; PF00520; Ion_trans; 1.
DR PRINTS; PR01097; TRANSRECEPTR.
DR SMART; SM00248; ANK; 2.
DR TIGRFAMs; TIGR00870; trp; 1.
DR PROSITE; PS50088; ANK_REPEAT; 1.
DR PROSITE; PS50297; ANK_REPEAT; 1.
KW Ionic channel; Transmembrane; Ion transport; Calcium channel;
KW ANK repeat; Repeat; Alternative splicing.
FT DOMAIN 1 331 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 332 352 POTENTIAL.
FT DOMAIN 353 364 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 365 385 POTENTIAL.
FT DOMAIN 386 438 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 439 459 POTENTIAL.
FT DOMAIN 460 471 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 472 492 POTENTIAL.
FT DOMAIN 493 513 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 514 534 POTENTIAL.
FT DOMAIN 535 601 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 602 622 POTENTIAL.
FT DOMAIN 623 981 CYTOPLASMIC (POTENTIAL).
FT REPEAT 71 100 ANK 1.
FT REPEAT 143 172 ANK 2.
FT DOMAIN 617 981 BINDS TO TRP1, TRP2 AND TRP3 (BY
FT SITE 979 981 SIMILARITY).
FT SITE 979 981 ESSENTIAL FOR BINDING TO NHERF PDZ DOMAIN
FT VARSPLIC 1 513 MISSING (IN ISOFORM BETA).
SQ SEQUENCE 981 AA; 112531 MW; 57B172FD65B791C9 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 981;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGP 5
DB 502 SHLGP 506

RESULT 196
A60D_DROME
ID A60D_DROME STANDARD; PRT; 1013 AA.
AC P91927; G9W160;
DT 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calcium-binding mitochondrial protein Anon-60Da.
GN ANON-60DA OR CG4589.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RN SEQUENCE FROM N.A.

```

EMBL; AF421359; AAI24550.1; -
 DR EMBL; U40983; AAC50630.1; -
 DR Genew; HGNC:12336; TRPC4.
 DR MIM; 603651; -
 DR InterPro; IPR002110; ANK.
 DR InterPro; IPR002111; Cat_channel_TrpL.
 DR InterPro; IPR000636; M+channel_n19.
 DR InterPro; IPR002153; Trans_recep.
 DR InterPro; IPR004729; Trp_CaChannel.
 DR Pfam; PF00023; ank; 2.
 DR Pfam; PF00520; Ion_trans; 1.
 DR PRINTS; PR01097; TRANSRECEPTR.
 DR SMART; SM00248; ANK; 2.
 DR TIGRfams; TIGR00870; trp; 1.
 DR PROSITE; PS50088; ANK_REPEAT; 1.
 DR PROSITE; PS50297; ANK_REP_REGION; 1.
 DR Ionic channel; Transmembrane; Ion transport; Calcium channel;
 KW ANK repeat; Repeat; Alternative splicing.
 FT DOMAIN 1 329
 FT TRANSMEM 330 350
 FT TRANSMEM 351 362
 FT TRANSMEM 363 383
 FT TRANSMEM 384 436
 FT TRANSMEM 437 457
 FT TRANSMEM 458 469
 FT TRANSMEM 470 490
 FT TRANSMEM 491 511
 FT TRANSMEM 512 532
 FT TRANSMEM 533 599
 FT TRANSMEM 600 620
 FT TRANSMEM 621 977
 FT DOMAIN 621 977
 FT REPEAT 69 98
 FT REPEAT 141 170
 FT DOMAIN 377 382
 FT SITE 615 977
 FT SITE 972 974
 FT VARSPIC 629 693
 FT VARSPIC 730 870
 FT VARSPIC 785 868
 FT SEQUENCE 977 AA; 112100 MW; 77E4D27C374D660E CRC64;
 Query Match 33.3%; Score 5; DB 1; Length 977;
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SHLP 5
 Db 500 SHLP 504
 RESULT 194
 TRP4_RAT
 ID TRP4_RAT STANDARD; PRT; 977 AA.
 AC 035119; Q9EQ75; Q9EQ74;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Short transient receptor potential channel 4 (TrpC4) (Trp4)
 DE (Capacitative calcium entry channel 1) (CCE1).
 ID TRP4.
 GN TRPC4.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus;
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM ALPHA).
 RC STRAIN=Mistar Itanichi; TISSUE=Brain;
 RX MEDLINE=97189270; PubMed=9037541;
 RA Funayama M., Goto K., Kondo H.;
 RT "Cloning and expression localization of cDNA for rat homolog of TRP
 protein, a possible store-operated calcium (Ca2+) channel.";

RL Brain Res. Mol. Brain Res. 43:259-266(1996).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS ALPHA AND BETA).
 RC STRAIN=Sprague-Dawley;
 RX MEDLINE=20545496; PubMed=10980202;
 RA Tang Y., Tang J., Chen Z., Trost C., Flocke V., Li M., Ramesh V.,
 RA Zhu M.X.;
 RT "Association of mammalian trp4 and phospholipase C isozymes with a PDZ
 RT domain-containing protein, NHERF.";
 RT J. Biol. Chem. 275:37559-37564(2000).
 CC -1- FUNCTION: THOUGHT TO FORM NON-SELECTIVE A RECEPTOR-ACTIVATED
 CC CALCIUM PERMEANT CATION CHANNEL. PROBABLY IS OPERATED BY A
 CC PHOSPHATIDYLINOSITOL SECOND MESSENGER SYSTEM ACTIVATED BY RECEPTOR
 CC TYROSINE KINASES OR G-PROTEIN COUPLED RECEPTORS. HAS ALSO BEEN
 CC SHOWN TO BE CALCIUM-SELECTIVE (BY SIMILARITY). MAY ALSO BE
 CC ACTIVATED BY INTRACELLULAR CALCIUM STORE DEPLETION.
 CC -1- SUBUNIT: ISOFORM ALPHA BUT NOT ISOFORM BETA ASSOCIATES WITH
 CC INOSITOL-1,4,5-TRIPHOSPHATE RECEPTOR (ITPR). INTERACTS WITH NHERF
 CC (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: ALPHA (SHOWN HERE) AND BETA: ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: BELONGS TO THE TRANSIENT RECEPTOR FAMILY. STRPC
 CC SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 2 ANK REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; AB008889; BAA23599.1; -
 DR EMBL; AF288407; AAG21809.1; -
 DR EMBL; AF288408; AAG21810.1; -
 DR InterPro: IPR002110; ANK.
 DR InterPro: IPR002111; Cat_channel_TrpL.
 DR InterPro: IPR000636; M+channel_n19.
 DR InterPro: IPR002153; Trans_recep.
 DR InterPro: IPR004729; Trp_CaChannel.
 DR Pfam; PF00023; ank; 2.
 DR Pfam; PF00520; Ion_trans; 1.
 DR PRINTS; PR01097; TRANSRECEPTR.
 DR SMART; SM00248; ANK; 2.
 DR TIGRfams; TIGR00870; trp; 1.
 DR PROSITE; PS50088; ANK_REPEAT; 1.
 DR PROSITE; PS50297; ANK_REP_REGION; 1.
 DR Ionic channel; Transmembrane; Ion transport; Calcium channel;
 KW ANK repeat; Repeat; Alternative splicing.
 FT DOMAIN 1 329
 FT TRANSMEM 330 350
 FT TRANSMEM 351 362
 FT TRANSMEM 363 383
 FT TRANSMEM 384 436
 FT TRANSMEM 437 457
 FT TRANSMEM 458 469
 FT TRANSMEM 470 490
 FT TRANSMEM 491 511
 FT TRANSMEM 512 532
 FT TRANSMEM 533 599
 FT TRANSMEM 600 620
 FT TRANSMEM 621 974
 FT REPEAT 69 98
 FT REPEAT 141 170
 FT DOMAIN 377 382
 FT SITE 615 977
 FT SITE 975 977
 FT VARSPIC 784 867
 FT VARSPIC 11 11
 FT CONFLICT 72 76
 FT CONFLICT 72 76
 BINDS TO ITPR1, ITPR2 AND ITPR3 (BY
 SIMILARITY).
 ESSENTIAL FOR BINDING TO NHERF PDZ DOMAIN
 (BY SIMILARITY).
 MISSING (IN ISOFORM BETA).
 N -> D (IN REF. 1).
 TALLI -> GLIFSL (IN REF. 1).

```

CC INTRACELLULAR CALCIUM STORE DEPLETION.
CC -1- SUBUNIT: INTERACTS WITH NHERF (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN. VERY LOW LEVELS DETECTED
CC IN LIVER KIDNEY, TESTIS, AND UTERUS.
CC -1- SIMILARITY: BELONGS TO THE TRANSIENT RECEPTOR FAMILY. STRPC
CC SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 2 ANK REPEATS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: AF029983; AAC13550.1; -.
DR EMBL: AF060107; AAF02200.1; -.
DR EMBL: AJ006204; CAA06912.1; -.
DR MGI: MGI:109524; Trp5.
DR InterPro: IPR002110; ANK.
DR InterPro: IPR002111; Cal_channel_TripL.
DR InterPro: IPR000636; M+channel_hlg.
DR InterPro: IPR002153; Trans_recep.
DR InterPro: IPR004729; Trp_CaChannel.
DR Pfam: PF00023; ank; 2.
DR Pfam: PF00520; Ion_trans; 1.
DR PRINTS: PR01097; TRNSRCEPTRP.
DR SMART: SM00248; ANK; 2.
DR TIGRFAMs: TIGR00870; trip; 1.
DR PROSITE: PS50088; ANK_REPEAT; FALSE_NEG.
DR PROSITE: PS50297; ANK_REP_REGION; FALSE_NEG.
DR Ionic channel: Transmembrane: Ion transport; Calcium channel;
KW ANK repeat; Repeat; Glycoprotein.
FT DOMAIN 1 330 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 331 351 POTENTIAL.
FT DOMAIN 352 398 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 399 419 POTENTIAL.
FT DOMAIN 420 437 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 438 458 POTENTIAL.
FT DOMAIN 459 470 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 471 491 POTENTIAL.
FT DOMAIN 492 512 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 513 533 POTENTIAL.
FT DOMAIN 534 603 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 604 624 POTENTIAL.
FT DOMAIN 625 975 POTENTIAL.
FT REPEAT 69 98 ANK 1.
FT REPEAT 141 170 POLY-ARG.
FT DOMAIN 690 693 ESSENTIAL FOR BINDING TO NHERF PDZ DOMAIN
FT SITE 971 973 (BY SIMILARITY).
FT CABOBYD 461 461 N-LINKED (GLNAC. .) (POTENTIAL).
FT SEQUENCE 975 AA: 111457 MW: 24816853.202 CR64;
Query Match 33.3%; Score 5; DB 1; Length 975;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHLCP 5
Db 501 SHLCP 505

```

```

DE protein 4) (htrp-4) (htrp4).
GN TRPC4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RC SEQUENCE FROM N.A. (ISOFORM ALPHA).
RC TISSUE=Kidney;
RX McKay R.R., Szymczek-Seay C.L., Lievremon J.-P., Bird G.S., Zilt C.,
RX Jiongling E., Lueckhoff A., Putney J.W. Jr.;
RT "Cloning and expression of the human transient receptor potential 4
RT (TRP4) gene: localization and functional expression of human TRP4 and
RT TRP3."
RL Biochem. J. 351:735-746(2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS ALPHA; BETA; DELTA AND GAMMA).
RP TISSUE=Embryonic kidney;
RX MEDLINE=21099836; PubMed=11163362;
RA Merry L., Magnifico F., Schmidt K., Krause K.-H., Dufour J.-F.;
RT "Alternative splice variants of htrp4 differentially interact with the
RT C-terminal portion of the inositol 1,4,5-trisphosphate receptors."
RL FEBS Lett. 487:377-383(2001).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS ALPHA AND BETA).
RX MEDLINE=21671347; PubMed=11713258;
RA Schaefer M., Plant T.D., Stresow N., Albrecht N., Schultz G.;
RT "Functional differences between TRPC4 splice variants."
RL J. Biol. Chem. 277:3752-3759(2002).
RN [4]
RP SEQUENCE OF 514-633 FROM N.A.
RP TISSUE=Kidney;
RX MEDLINE=96234226; PubMed=8646775;
RA Zhu X., Jiang M., Peyton M., Boulay G., Hurst R., Stefani E.,
RA Birdamer L.;
RT "trp, a novel mammalian gene family essential for agonist-activated
RT capacitative Ca2+ entry."
RL Cell 85:661-671(1996).
CC -1- FUNCTION: THOUGHT TO FORM A RECEPTOR-ACTIVATED NON-SELECTIVE
CC CALCIUM PERMEANT CATION CHANNEL. PROBABLY IS OPERATED BY A
CC PHOSPHATIDYLINOSITOL SECOND MESSENGER SYSTEM ACTIVATED BY RECEPTOR
CC TYROSINE KINASES OR G-PROTEIN COUPLED RECEPTORS. HAS ALSO BEEN
CC SHOWN TO BE CALCIUM-SELECTIVE (BY SIMILARITY). MAY ALSO BE
CC ACTIVATED BY INTRACELLULAR CALCIUM STORE DEPLETION.
CC -1- SUBUNIT: ISOFORM ALPHA BUT ISOFORM BETA ASSOCIATES WITH INOSITOL-
CC 1,4,5-TRIPHOSPHATE RECEPTOR (ITPR). INTERACTS WITH NHERF (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -1- ALTERNATIVE PRODUCTS: 4 ISOFORMS: ALPHA (SHOWN HERE), BETA, DELTA
CC AND GAMMA; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN PLACENTA. EXPRESSED AT
CC LOWER LEVELS IN HEART, PANCREAS, KIDNEY AND BRAIN. ISOFORM ALPHA
CC WAS FOUND TO BE THE PREDOMINANT ISOFORM. ISOFORM BETA WAS NOT
CC FOUND IN PANCREAS AND BRAIN.
CC -1- SIMILARITY: BELONGS TO THE TRANSIENT RECEPTOR FAMILY. STRPC
CC SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 2 ANK REPEATS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: AF063822; AAF22927.1; -.
DR EMBL: AF063823; AAF22928.1; -.
DR EMBL: AF063824; AAF22929.1; -.
DR EMBL: AF063825; AAF22930.1; -.
DR EMBL: AF175406; AAD51736.1; -.
DR EMBL: AF421358; AAL24549.1; -.

```

ID TRP5_RABIT STANDARD: PRT; 974 AA.
AC 062852;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Short transient receptor potential channel 5 (trp5) (Rtrp5)
DE (Capacitative calcium entry channel 2) (CCE2).
GN TRP5 OR TRP5.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_Taxid=9986;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98353453; PubMed=9687496;
RA Philipp S., Hambrecht J., Braslavski L., Schroth G., Freichel M.,
RT Murakami M., Cavalle A., Flockerzi V.;
RT "A novel capacitative calcium entry channel expressed in excitable
cells";
RT EMBO J. 17:4274-4282(1998).
CC -1- FUNCTION: THOUGHT TO FORM A RECEPTOR-ACTIVATED NON-SELECTIVE
CC CALCIUM PERMEANT CATION CHANNEL. PROBABLY IS OPERATED BY A RECEPTOR
CC PHOSPHATIDYLINOSITOL SECOND MESSENGER SYSTEM ACTIVATED BY RECEPTOR
CC TيروسINE KINASES OR G-PROTEIN COUPLED RECEPTORS. HAS ALSO BEEN
CC SHOWN TO BE CALCIUM-SELECTIVE. MAY ALSO BE ACTIVATED BY
CC INTRACELLULAR CALCIUM STORE DEPLETION.
CC -1- SUBUNIT: INTERACTS WITH NHERF (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN BRAIN.
CC -1- SIMILARITY: BELONGS TO THE TRANSIENT RECEPTOR FAMILY. STRPC
CC SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 2 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AJ006203; CA006911.1; -.
DR InterPro: IPR002110; ANK.
DR InterPro: IPR002111; Cat_channel_TrpL.
DR InterPro: IPR000636; M+channel_nlg.
DR InterPro: IPR002153; Trans_recep.
DR InterPro: IPR004729; Trp_CaChannel.
DR Pfam: PF00023; ank; 2.
DR Pfam: PF00520; Ion_trans; 1.
DR PRINTS: PR01097; TRANSRECEPTR.
DR SMART: SM00248; ANK; 2.
DR TIGRFS: TIGR00870; trp; 1.
DR PROSITE: PS50088; ANK_REPEAT; FALSE_NEG.
DR PROSITE: PS50297; ANK_REP_REGION; FALSE_NEG.
KM Ionic channel; Transmembrane; Ion transport; Calcium channel;
KM ANK repeat; Repeat; Glycoprotein.
FT DOMAIN 1 330 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 331 351 POTENTIAL.
FT DOMAIN 352 398 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 399 419 POTENTIAL.
FT DOMAIN 420 437 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 438 458 POTENTIAL.
FT DOMAIN 459 470 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 471 491 POTENTIAL.
FT DOMAIN 492 512 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 513 533 POTENTIAL.
FT DOMAIN 534 603 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 604 624 POTENTIAL.
FT DOMAIN 625 974 CYTOPLASMIC (POTENTIAL).
FT REPEAT 69 98 ANK 1.
FT REPEAT 141 170 ANK 2.
FT DOMAIN 690 693 POLY-ARG.

FT SITE 972 974 ESSENTIAL FOR BINDING TO NHERF PDZ DOMAIN
FT CAROHD 461 461 (BY SIMILARITY).
SQ SEQUENCE 974 AA: 111536 MW: 94255E0F2B925316 CRC64:
Query Match 33.3%; Score 5; DB 1; Length 974;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHLGP 5
Db 501 SHLGP 505
RESULT 192
TRP5_MOUSE
ID TRP5_MOUSE STANDARD: PRT; 975 AA.
AC Q9QX29; Q9R0D4; Q9QWT1; Q61059;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Short transient receptor potential channel 5 (trp5) (Transient
DE receptor protein 5) (Mtrp5) (trp-related protein 5) (Capacitative
DE calcium entry channel 2) (CCE2).
GN TRP5 OR TRP5 OR TRP5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98221157; PubMed=9553080;
RA Okada T., Shimizu S., Makamori M., Maeda A., Kurosaki T., Takada N.,
RT Imoto K., Mori Y.;
RT "Molecular cloning and functional characterization of a novel
RT receptor-activated TRP Ca2+ channel from mouse brain.";
RT J. Biol. Chem. 273:10279-10287(1998).
RN (2)
RP SEQUENCE FROM N.A.
RA Zhu X., Peyton M., Boulay B., Birnbaumer B.;
RT "Molecular cloning and functional expression of mouse TRP5";
RT Submitted (Apr-1998) to the EMBL/GenBank/DBJ databases.
RN (3)
RP SEQUENCE FROM N.A.
RX MEDLINE=20545496; PubMed=10980202;
RA Tang Y., Tang J., Chen Z., Trost C., Flockerzi V., Li M., Ramesh V.,
RT Zhu M.X.;
RT "Association of mammalian trp4 and phospholipase C isozymes with a PDZ
RT domain-containing protein, NHERF";
RT J. Biol. Chem. 275:37559-37564(2000).
RN (4)
RP SEQUENCE OF 1-966 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98353453; PubMed=9687496;
RA Philipp S., Hambrecht J., Braslavski L., Schroth G., Freichel M.,
RT Murakami M., Cavalle A., Flockerzi V.;
RT "A novel capacitative calcium entry channel expressed in excitable
RT cells";
RT EMBO J. 17:4274-4282(1998).
RN (5)
RP SEQUENCE OF 515-637 FROM N.A.
RX MEDLINE=96234226; PubMed=8646775;
RA Zhu X., Jiang M., Peyton M., Boulay G., Hurst R., Stefani F.,
RT Birnbaumer B.;
RT "trp, a novel mammalian gene family essential for agonist-activated
RT capacitative Ca2+ entry";
RT Cell 85:661-671(1996).
CC -1- FUNCTION: THOUGHT TO FORM A RECEPTOR-ACTIVATED NON-SELECTIVE
CC CALCIUM PERMEANT CATION CHANNEL. PROBABLY IS OPERATED BY A
CC PHOSPHATIDYLINOSITOL SECOND MESSENGER SYSTEM ACTIVATED BY RECEPTOR
CC TYROSINE KINASES OR G-PROTEIN COUPLED RECEPTORS. HAS ALSO BEEN
CC SHOWN TO BE CALCIUM-SELECTIVE. MAY ALSO BE ACTIVATED BY

RESULT 190
ID TRP4_MOUSE STANDARD: 062350; PRT: 974 AA.
AC 090005: 090009: 062350;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Short transient receptor potential channel 4 (TRPC4) (Receptor-
activated cation channel TRP4) (Capacitative calcium entry channel
TRP4) OR TRP4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS ALPHA AND BETA).
RC TISSUE-Brain;
RA Zhu X., Boulay G., Jiang M., Birnbaumer L.;
RT "Trp4 is involved in capacitative calcium entry in murine cells.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS ALPHA AND BETA).
RA Qian F., Philipson L.H.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS ALPHA AND BETA).
RC TISSUE-Brain;
RA Mori Y., Takada N., Okada T., Wakamori M., Imoto K., Wanifuchi H.,
RA Oka H., Oba A., Ikenaka K., Kuroseki T.;
RT "Differential distribution of TRP Ca2+ channel isoforms in mouse
brain.";
RL NeuroReport 9:507-515(1998).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORMS ALPHA AND BETA).
RA MEDLINE=20545496; PubMed=10980202;
RT Tang Y., Tang J., Chen Z., Trost C., Flockerzi V., Li M., Ramesh V.,
RA Zhu M.X.;
RT "Association of mammalian trp4 and phospholipase C isozymes with a p2z
domain-containing protein, NHERF.";
RL J. Biol. Chem. 273:3759-3764(2000).
RN [5]
RP SEQUENCE OF 505-642 FROM N.A.
RC TISSUE-Brain;
RA MEDLINE=96003894; PubMed=7575478;
RT Petersen C.C.H., Berridge M.J., Borge M.F., Bennett D.L.;
RT "Putative capacitative calcium entry channels: expression of
Drosophila trp and evidence for the existence of vertebrate
homologues.";
RL Biochem. J. 311:41-44(1995).
RN [6]
RP FUNCTION
RA MEDLINE=21113116; PubMed=11175743;
RA Freichel M., Suh S.H., Pfeifer A., Schweig U., Trost C.,
RA Weisgerber P., Biel M., Philipp S., Freise D., Droogmans G.,
RA Hofmann F., Flockerzi V., Nilius B.;
RT "Lack of an endothelial store-operated Ca2+ current impairs
agonist-dependent vasorelaxation in TRP4-/- mice.";
RL Nat. Cell Biol. 3:121-127(2001).
CC -1- FUNCTION: THOUGHT TO FORM A RECEPTOR-ACTIVATED NON-SELECTIVE
CALCIUM PERMEANT CATION CHANNEL. PROBABLY IS OPERATED BY A
PHOSPHATIDYLINOSITOL SECOND MESSENGER SYSTEM ACTIVATED BY RECEPTOR
TYROSINE KINASES OR G-PROTEIN COUPLED RECEPTORS. HAS ALSO BEEN
CC SHOWN TO BE CALCIUM-SELECTIVE (BY SIMILARITY). MAY ALSO BE
CC ACTIVATED BY INTRACELLULAR CALCIUM STORE DEPLETION. TRPC4
CC DEFICIENT MICE LACK A STORE-OPERATED CALCIUM ENTRY IN ENDOTHELIAL
CC CELLS.
CC -1- SUBUNIT: ISOFORM ALPHA BUT NOT ISOFORM BETA ASSOCIATES WITH
CC INOSITOL-1,4,5-TRIPHOSPHATE RECEPTOR (ITPR) (BY SIMILARITY).
CC INTERACTS WITH NHERF.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: ALPHA (SHOWN HERE) AND BETA. ARE

CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: ABUNDANTLY EXPRESSED IN BRAIN (HIPPOCAMPAL CA1
CC PYRAMIDAL NEURONS, DENTATE GYRUS GRANULE CELLS, AND CEREHAL
CC CORTICAL NEURONS, AND IN THE SEPTAL NUCLEI AND THE MITRAL LAYER OF
CC OLFACTORY BULB). LOWER LEVELS ARE DETECTED IN OTHER TISSUES.
CC -1- SIMILARITY: BELONGS TO THE TRANSIENT RECEPTOR FAMILY. STRC
CC SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 2 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF011543; AAD10167.1; -;
CC EMBL: U50922; AAC05179.1; -;
CC EMBL: AF190646; AAF01469.1; -;
CC EMBL: U50921; AAC05178.1; -;
CC EMBL: AF019663; AAD10168.1; -;
CC EMBL: X90697; CAA62230.1; -;
CC MGD: MGI:109525; Trp4.
CC InterPro: IPR002110; ANK.
CC InterPro: IPR002111; Cal_channel_TrpL.
CC InterPro: IPR000635; M_channel_nlg.
CC InterPro: IPR002133; Trans_recep.
CC InterPro: IPR004729; Trp_CaChannel.
CC Pfam: PF00023; ank. 2.
CC Pfam: PF00520; Ion_trans. 1.
CC PRINTS: PR01097; TRNSRECEPT.
CC SMART: SM00248; ANK. 2.
CC TRGFAMS; TIGR00870; trp. 1.
CC PROSITE; PS50088; ANK_REPEAT. 1.
CC PROSITE; PS50297; ANK_REPEAT_REGION. 1.
CC KW Ionic channel; Transmembrane; Ion transport; Calcium channel;
CC ANK repeat; Repeat; Alternative splicing.
CC FT DOMAIN 1 329
CC FT TRANSMEM 330 350
CC FT DOMAIN 351 362
CC FT TRANSMEM 351 362
CC FT TRANSMEM 363 383
CC FT TRANSMEM 384 436
CC FT TRANSMEM 437 457
CC FT TRANSMEM 458 469
CC FT TRANSMEM 470 490
CC FT DOMAIN 491 511
CC FT TRANSMEM 512 532
CC FT TRANSMEM 533 599
CC FT TRANSMEM 600 620
CC FT TRANSMEM 621 974
CC FT REPEAT 69 98
CC FT REPEAT 141 170
CC FT REPEAT 141 170
CC FT DOMAIN 615 974
CC FT SITE 972 974
CC FT VARSPIC 781 864
CC FT CONFLICT 780 780
CC FT CONFLICT 890 890
CC FT SEQUENCE 974 AA; 11574 MW; 2D0BB2F235F8D1 CRC64;
CC
CC Query Match 33.3%; Score 5; DB 1; Length 974;
CC Best Local Similarity 100.0%; Pred. No. 3 6e+02;
CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC RESULT 191
CC TRP5_RABBIT

```

RN [12]
RP SEQUENCE FROM N.A.
RA Andre B., Iragui Houssaini I., Urrestarazu L.A., Vissers S.;
RL Submitted (MAY-1996) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: TO DROSOPHILA SHUTTLE CRAFT PROTEIN (STC) AND HUMAN
CC NFX1.
CC
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
CC -----
CC [This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).]
CC
DR EMBL: 271299; CAA95885.1; -
DR SGD: S0004968; YNL023C.
DR InterPro: IPR001374; R3H.
DR InterPro: IPR000967; Znf_NFX1.
DR InterPro: IPR001841; Znf_ring.
DR Pfam: PF01422; zf-NF-X1; 5.
DR Pfam: PF01424; R3H; 1.
DR SMART: SM00393; R3H; 1.
DR SMART: SM00184; RING; 1.
DR SMART: SM00438; Znf_NFX; 7.
DR PROSITE: PS00518; ZF_RING_1; FALSE_NEG.
DR PROSITE: PS00089; ZF_RING_2; 1.
DR Hypothetical protein: ZINC-finger; Repeat.
KW ZN_RING 68 118 RING-TYPE.
FT FT 150 644 7 X APPROXIMATE REPEATS, CYS-RICH.
FT REPEAT 150 185 1.
FT REPEAT 206 251 2.
FT REPEAT 273 330 3.
FT REPEAT 352 390 4.
FT REPEAT 458 497 5.
FT REPEAT 575 610 6.
FT REPEAT 611 644 7.
SQ SEQUENCE 965 AA; 108494 MW; 121C57BB0766FA9D CRC64;
OY 2 HLGP 6
DB 168 HLGP 172
PRT; 973 AA.
ID TRP5_HUMAN STANDARD;
DC Q9UL62; O75233; Q9Y514;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Short transient receptor potential channel 5 (TrpC5) (Htrp5).
DE TRPC5 OR TRP5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN [12]
RP SEQUENCE FROM N.A.
RP TISSUE=fetal brain;
RX MEDLINE=99425273; PubMed=10493832;
RX Sossey-Alaoui K., Lyon J.A., Jones L., Abidi F.E., Hartung A.J.,
RX Haase B., Schwartz C.E., Stevenson R.E., Srivastava A.K.;
RT "Molecular cloning and characterization of TRPC5 (HTRP5), the human
RT homolog of a mouse brain receptor-activated capacitative Ca(2+)
RT entry channel."
RL Genomics 60:330-340(1999).
[2]

```

```

RP SEQUENCE FROM N A.
RA Rohlfing T., Langston Y., Mead K., Bird C.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: THOUGHT TO FORM A RECEPTOR-ACTIVATED NON-SELECTIVE
CC CALCIUM PERMANENT CATION CHANNEL. PROBABLY IS OPERATED BY A
CC PHOSPHATIDYLINOSITOL SECOND MESSENGER SYSTEM ACTIVATED BY RECEPTOR
CC TYROSINE KINASES OR G-PROTEIN COUPLED RECEPTORS. HAS ALSO BEEN
CC SHOWN TO BE CALCIUM-SELECTIVE (BY SIMILARITY). MAY ALSO BE
CC ACTIVATED BY INTRACELLULAR CALCIUM STORE DEPLETION.
CC -I- SUBUNIT: INTERACTS WITH NHERF (BY SIMILARITY).
CC -I- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -I- TISSUE SPECIFICITY: EXPRESSED IN BRAIN WITH HIGHER LEVELS IN FETAL
CC BRAIN.
CC -I- SIMILARITY: BELONGS TO THE TRANSIENT RECEPTOR FAMILY. STRPC
CC SUPERFAMILY.
CC -I- SIMILARITY: CONTAINS 2 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
CC EMBL; AF054568; AAF00002.1; -
CC EMBL; AC005191; AAC24565.1; -
CC EMBL; AL049563; CAB44737.1; -
CC Genew; HGNC:12337; TRPC5.
CC MIM; 300334; -
CC DR InterPro; IPR002110; ANK.
CC DR InterPro; IPR002111; Cat_channel_Trfp.
CC DR InterPro; IPR000636; M-channel_nlg.
CC DR InterPro; IPR002153; Trans_recep.
CC DR InterPro; IPR004729; Trp_Cachannel.
CC Pfam; PF00023; ank; 2.
CC Pfam; PF00520; Ion_trans; 1.
CC DR PRINTS; PR01097; TRANSCREPTPR.
CC DR SMART; SM00248; ANK; 2.
CC DR TIGRFAMS; TIGR00870; trp; 1.
CC DR PROSITE; PS50088; ANK_REPEAT; FALSE_NEG.
CC DR PROSITE; PS50297; ANK_REP_REGION; FALSE_NEG.
CC KW Ionic channel; Transmembrane; Ion transport; Calcium channel;
CC ANK repeat; Repeat; Glycoprotein.
CC FT DOMAIN 1 330
CC FT TRANSMEM 331 351 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 352 398 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 399 419 POTENTIAL.
CC FT DOMAIN 420 437 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 438 458 POTENTIAL.
CC FT DOMAIN 459 470 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 471 491 POTENTIAL.
CC FT DOMAIN 492 512 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 513 533 POTENTIAL.
CC FT DOMAIN 534 603 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 604 624 POTENTIAL.
CC FT DOMAIN 625 973 CYTOPLASMIC (POTENTIAL).
CC FT REPEAT 69 98 ANK 1.
CC FT REPEAT 141 170 ANK 2.
CC FT SITE 971 973
CC FT CARBOHYD 461 461
CC FT SEQUENCE 973 AA; 111411 MW; FBC8CBF17BE42166 CRC64;
CC
CC Query Match 33.3%; Score 5; DB 1; Length 973;
CC Best Local Similarity 100.0%; Pred. No. 3.6e+02;
CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC 1 SHLP 5
CC |||||
CC 501 SHLP 505

```

Db 719 PESRA 723

RESULT 186

UVRA_PASMU STANDARD: PRT: 943 AA.

AC P53979;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Excinuclease ABC subunit A.

GN UVRA OR PM1951.

OS Pasteurella multocida.

OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;

OC Pasteurella.

OX NCBI_TaxID=747;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=pm70;

RX MEDLINE=21145866; PubMed=11248100;

RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.,

"Complete genomic sequence of Pasteurella multocida pm70.";

RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).

CC -1- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). UVRA IS AN ATPASE AND A DNA-BINDING PROTEIN THAT PREFERENTIALLY BINDS SINGLE-STRANDED OR UV-IRRADIATED DOUBLE-STRANDED DNA (BY SIMILARITY).

CC -1- SUBUNIT: CONSISTS OF THREE SUBUNITS: UVRA, UVRA AND UVRC.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).

CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. UVRA SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AE006231; AK04035.1; -

DR InterPro: IPR003439; ABC_transportr.

DR InterPro: IPR004602; UVRA.

DR Pfam: PF00005; ABC_tran; 2.

DR ProDom: PD000006; ABC_transportr; 1.

DR TIGRFAMs: TIGR00630; uvra; 1.

DR PROSITE: PS00211; ABC_TRANSPORTER; 2.

KW SOS response; Excision nuclease; DNA repair; ATP-binding; Repeat; DNA-binding; Zinc-finger; Complete proteome.

KW NP_BIND 31 38 ATP (POTENTIAL).

FT NP_BIND 640 647 ATP (POTENTIAL).

FT ZN_FING 253 280 C4-TYPE.

FT ZN_FING 740 766 C4-TYPE.

SEQUENCE 943 AA; 104186 MW; A096DB1162D3C354 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 943;

Best Local Similarity 100.0%; Pred. No. 3.5e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRA 14

Db 719 PESRA 723

RESULT 187

KINL_LEICH STANDARD: PRT: 955 AA.

AC P46865;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE Kinesin-like protein K39 (fragment).

GN KIN.

OS Leishmania chagasi.

OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.

OX NCBI_TaxID=44271;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=HOW/BR/82 / Isolate BR-2;

RX MEDLINE=93133867; PubMed=8421715;

RA Burns J.M. Jr., Shreffler W.G., Benson D.R., Ghailb H.W., Badaro R.,

RA Reed S.G.;

RT "Molecular characterization of a kinesin-related antigen of Leishmania chagasi that detects specific antibody in African and American visceral leishmaniasis.";

RL Proc. Natl. Acad. Sci. U.S.A. 90:775-779(1993).

CC -1- DEVELOPMENTAL STAGE: PREDOMINANT IN AMASTIGOTES.

CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: L07879; AA29254.1; -

DR HSSP: P17119; 3KAR.

DR InterPro: IPR001752; Kinesin_motor.

DR Pfam: PF00225; kinesin; 1.

DR SMART: SM00129; Kisc; 1.

DR PROSITE: PS00411; KINESIN_MOTOR_DOMAIN1; 1.

DR PROSITE: PS50067; KINESIN_MOTOR_DOMAIN2; 1.

KW Motor protein; Microtubules; ATP-binding; Coiled coil; Repeat.

KW DOMAIN 1 399 KINESIN_MOTOR (BY SIMILARITY).

FT NP_BIND 122 129 COILED COIL (POTENTIAL).

FT DOMAIN 704 742 ATP (POTENTIAL).

FT REPEAT 704 742 1.

FT REPEAT 743 781 2.

FT REPEAT 782 820 3.

FT REPEAT 821 859 4.

FT REPEAT 860 898 5.

FT REPEAT 899 937 6.

FT REPEAT 938 955 7 (PARTIAL).

FT NON_TER 955 955

SEQUENCE 955 AA; 106168 MW; 8CA76815BE84C6B9 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 955;

Best Local Similarity 100.0%; Pred. No. 3.5e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRAA 15

Db 478 ESRAA 482

RESULT 188

YNC3_YEAST STANDARD: PRT: 965 AA.

ID YNC3_YEAST

AC P53971;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Hypothetical 108.5 kDa protein in UMG3-HDA1 intergenic region.

GN YNL023C OR N2812.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

OX NCBI_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RA Duesterhoeft A., Floeth M., Fritz C., Heuss-Neitzel D.,

RA Hilbert H., Moestl D.;

RL submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.

```
CC -1- FUNCTION: The ABC excision nuclease is a DNA repair enzyme that
CC catalyzes the excision reaction of UV-damaged nucleotide segments
CC producing oligomers having the modified base(s). UvrA is an ATPase
CC and a DNA-binding protein that preferentially binds single-
CC stranded or UV-irradiated double-stranded DNA (by similarity).
CC -1- SUBUNIT: Consists of three subunits; uvrA, uvrB and uvrC.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. UVRB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE005638; AAG59256.1; -.
DR EMBL; AP002568; BAB38463.1; -.
DR InterPro: IPR003439; ABC_transportr.
DR InterPro: IPR004602; UvrA.
DR Pfam: PF00005; ABC_tran; 2.
DR ProDom: PD000006; ABC_transportr; 1.
DR TIGRfams: TIGR00630; uvrA; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
KW SOS response; Excision nuclease; DNA repair; ATP-binding; Repeat;
KM DNA-binding; Zinc-finger; Complete proteome.
FT NP_BIND 31 38 ATP (POTENTIAL).
FT NP_BIND 640 647 ATP (POTENTIAL).
FT ZN_FING 253 280 C4-TYPE.
FT ZN_FING 740 766 C4-TYPE.
SQ SEQUENCE 940 AA; 103884 MW; A20C90C935A0ACB CRC64;

Query Match 33.3%; Score 5; DB 1; Length 940;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRA 14
Db 719 PESRA 723

RESULT 185
UVR_A_ECOLI STANDARD; PRT; 940 AA.
AC P07671; P76788;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Exclunuclease ABC subunit A.
GN UVR_A OR DINE OR B4058.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxId=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86168204; PubMed=3007478;
RA Huslin I., van Houten B., Thomas D.C., Sancar A.;
RT "Sequences of Escherichia coli uvrA gene and protein reveal two
RT potential ATP binding sites.";
RL J. Biol. Chem. 261:4895-4901(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / MG1655;
RX MEDLINE=94089392; PubMed=8265357;
RA Blattner F.R., Burland V.D., Plunkett G. III, Sofia H.J.,
RA Daniels D.L.;
RT "Analysis of the Escherichia coli genome. IV. DNA sequence of the
RT region from 89.2 to 92.8 minutes.";
RL Nucleic Acids Res. 21:5408-5417(1993).
RN [3]
RP SEQUENCE OF 1-25 FROM N.A.
```

```
RX MEDLINE=83299251; PubMed=6310514;
RA Kanderof C., Brandsma J.A., Kartasova T., van de Putte P.;
RT "In vivo regulation of the uvrA gene: role of the '-10' and '-35'
RT promoter regions.";
RL Nucleic Acids Res. 11:5795-5810(1983).
RN [4]
RP SEQUENCE OF 1-14 FROM N.A.
RX MEDLINE=82220077; PubMed=6283374;
RA Sancar A., Sancar G.B., Rupp M.D., Little J.W., Mount D.W.;
RT "LexA protein inhibits transcription of the E. coli uvrA gene in
RT vitro.";
RL Nature 298:96-98(1982).
RN [5]
RP CHARACTERIZATION.
RX MEDLINE=91208117; PubMed=1826851;
RA Myles G.M., Sancar A.;
RT "Isolation and characterization of functional domains of UvrA.";
RL Biochemistry 30:3834-3840(1991).
RN [6]
RP MUTAGENESIS OF CYS-253.
RX MEDLINE=89380205; PubMed=2550431;
RA Navaratnam S., Myles G.M., Strange R.W., Sancar A.;
RT "Evidence from extended X-ray absorption fine structure and site-
RT specific mutagenesis for zinc fingers in UvrA protein of Escherichia
RT coli.";
RL J. Biol. Chem. 264:16067-16071(1989).
CC -1- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT
CC CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS
CC PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). UVR A IS AN ATPASE
CC AND A DNA-BINDING PROTEIN THAT PREFERENTIALLY BINDS SINGLE-
CC STRANDED OR UV-IRRADIATED DOUBLE-STRANDED DNA.
CC -1- SUBUNIT: CONSISTS OF THREE SUBUNITS; UVR A, UVR B AND UVR C.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- MISCELLANEOUS: BINDS ABOUT 2 ZINC ATOMS/MOLECULE.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. UVRB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M13495; AAA24754.1; -.
DR EMBL; U00006; AAC43152.1; -.
DR EMBL; AE000479; AAC77028.1; -.
DR EMBL; X01621; CAA25764.1; -.
DR EMBL; J01721; AAA24753.1; -.
DR PIR; A23869; BVECUA.
DR ECODBASE; H124.0; 6TH EDITION.
DR Ecogene; E811061; uvrA.
DR InterPro: IPR003439; ABC_transportr.
DR InterPro: IPR004602; UvrA.
DR Pfam: PF00005; ABC_tran; 2.
DR ProDom: PD000006; ABC_transportr; 1.
DR TIGRfams: TIGR00630; uvrA; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
KW SOS response; Excision nuclease; DNA repair; ATP-binding; Repeat;
KM DNA-binding; Zinc-finger; Complete proteome.
FT NP_BIND 31 38 ATP.
FT NP_BIND 640 647 ATP.
FT ZN_FING 253 280 C4-TYPE.
FT ZN_FING 740 766 C4-TYPE.
FT MUTAGEN 253 253 C->A,H,S. REDUCED ACTIVITY.
SQ SEQUENCE 940 AA; 103867 MW; D61AAEB6514B60C CRC64;

Query Match 33.3%; Score 5; DB 1; Length 940;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRA 14
|||||
```

```

RL Mol. Biol. Cell 6:1605-1618(1995).
RN [2]
RP REVISIONS.
RA Krichner J.;
RL Submitted (MAR-1997) to the EMBL/Genbank/DBJ databases.
CC
CC -1- FUNCTION: KINESIN IS A MICROTUBULE-ASSOCIATED FORCE-PRODUCING
CC PROTEIN THAT MAY PLAY A ROLE IN ORGANELLE TRANSPORT. ITS MOTOR
CC ACTIVITY IS DIRECTED TOWARD THE MICROTUBULE'S PLUS END.
CC -1- DOMAIN: COMPOSED OF THREE STRUCTURAL DOMAINS: A LARGE GLOBULAR N-
CC TERMINAL DOMAIN WHICH IS RESPONSIBLE FOR THE MOTOR ACTIVITY OF
CC KINESIN (IT HYDROLYZES ATP AND BINDS MICROTUBULE), A CENTRAL
CC ALPHA-HELICAL COILED COIL DOMAIN THAT MEDIATES THE HEAVY CHAIN
CC DIMERIZATION, AND A SMALL GLOBULAR C-TERMINAL DOMAIN WHICH
CC INTERACTS WITH OTHER PROTEINS (SUCH AS THE KINESIN LIGHT CHAINS),
CC VESICLES AND MEMBRANOUS ORGANELLES.
CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L47106; AAB52961.1; -.
DR HSSP: P33176; 1B62.
DR InterPro: IPR001752; kinesin_motor.
DR Pfam: PF00225; kinesin_1.
DR PRINTS: PR00380; KINESINHEAVY.
DR SMART: SM00129; KISC; 1.
DR PROSITE: PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE: PS50067; KINESIN_MOTOR_DOMAIN2; 1.
DR Motor protein; Microtubules; ATP-binding; Coiled coil.
FT DOMAIN 1 337 KINESIN-MOTOR (BY SIMILARITY).
FT DOMAIN 338 741 GLOBULAR.
FT NP_BIND 742 928 COILED COIL.
FT NP_BIND 88 95 ATP (POTENTIAL).
FT NP_BIND 238 245 ATP (POTENTIAL).
SQ SEQUENCE 928 AA; 102410 MW; 2E2475195F674C02 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 928;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRA 14
DB 412 PESRA 416

RESULT 183
MSHA_HUMAN
ID MSHA_HUMAN STANDARD: PRT; 936 AA.
AC 015457;
DT 15-DEC-1998 (rel. 37, Created)
DT 15-DEC-1998 (rel. 37, Last sequence update)
DT 16-OCT-2001 (rel. 40, Last annotation update)
DE Muts protein homolog 4.
GN MSH4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCB1_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97446137; PubMed=9299235;
RA Paquis-Flickinger V., Santucci-Darmanin S., Paul R., Saunieres A.,
RA Turc-Carel C., Desnuelle C.;
RT "Cloning and expression analysis of a meiosis-specific Muts homolog:
RT the human MSH4 gene."
RT Genomics 44:188-194(1997).
CC -1- FUNCTION: INVOLVED IN MEIOTIC RECOMBINATION. REQUIRED FOR

```

```

CC RECIPROCAL RECOMBINATION AND PROPER SEGREGATION OF HOMOLOGOUS
CC CHROMOSOMES AT MEIOSIS.
CC -1- SUBUNIT: HETEROODIGOMER OF MSH4 AND MSH5 (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U89293; AAB72039.1; -.
DR Genew: HGNC:7327; MSH4.
DR MIM: 602105; -.
DR InterPro: IPR000432; Muts_C.
DR InterPro: IPR002863; Muts_N.
DR Pfam: PF00488; Muts_C; 1.
DR Pfam: PF01624; Muts_N; 1.
DR ProDom: PD001263; Muts_C; 1.
DR SMART: SM00534; Mutsac; 1.
DR SMART: SM00533; Mutsd; 1.
DR PROSITE: PS00486; DNA_MISMATCH_REPAIR_2; 1.
DR Meiosis; ATP-binding; DNA-binding.
FT NP_BIND 680 687 ATP (POTENTIAL).
SQ SEQUENCE 936 AA; 104774 MW; 5C5E91F5212D9D16 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 936;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPE 11
DB 877 RSTPE 881

RESULT 184
UVRA_ECO57
ID UVRA_ECO57 STANDARD: PRT; 940 AA.
AC 08X509;
DT 15-JUN-2002 (rel. 41, Created)
DT 15-JUN-2002 (rel. 41, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Excinuclease ABC subunit A.
GN UVRA OR Z5657 OR ECS5040.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCB1_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G., III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohemorrhagic Escherichia coli O157:H7."
RT Nature 409:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RT DNA Res. 8:11-22(2001).

```

MAP KINASE KINASE SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: U14636; AAA57280.1; -
DR HSSP: U23789; AAB17123.1; -
DR HSSP: P12931; IFMK.
DR MGD: MG1:134681; Map3K12.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR004040; STY_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00069; Pkinase.1.
DR ProDom: PD000001; Euk_pkinase.1.
DR SMART: SM00221; STYK; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation; Magnesium; Membrane.
FT DOMAIN 158 399
FT NP_BIND 164 172 ATP (BY SIMILARITY).
FT BINDING 185 185
FT ACT_SITE 269 269 BY SIMILARITY.
FT DOMAIN 56 62
FT DOMAIN 668 671
FT DOMAIN 753 758
FT DOMAIN 753 758
FT MUTAGEN 185 185 K->A: NO CATALYTIC ACTIVITY.
FT MUTAGEN 192 192 E->A: NO CHANGE.
FT CONFLICT 18 18 V->A (IN REF. 2).
FT CONFLICT 28 29 KL->N (IN REF. 2).
FT CONFLICT 382 382 S->P (IN REF. 2).
FT CONFLICT 494 495 EQ->DE (IN REF. 2).
FT CONFLICT 517 517 N->D (IN REF. 2).
FT CONFLICT 794 794 E->G (IN REF. 2).
SQ SEQUENCE 888 AA; 96083 MW; CFECDID34F889AB CRC64;

Query Match 33.3%; Score 5; DB 1; Length 888;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRA 14
| | | | |
Db 98 PESRA 102

RESULT 181
M3KC_RAT

ID	M3KC_RAT	STANDARD;	PRT;	888 AA.
AC	O63796;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Mitogen-activated protein kinase kinase kinase 12 (BC 2.7.1.37)			
DE	(MAPK-upstream kinase) (MOK).			
GN	MAP3K12 OR MOK.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96226099; PubMed=8637721;			
RA	Hirai S., Izawa M., Osada S., Sgyrou G., Ohno S.;			
RT	"Activation of the JNK pathway by distantly related protein kinases,			
RT	MKK and MOK."			
RL	Oncogene 12:641-650(1996).			

CC -1- FUNCTION: May be an activator of the JNK/SAPK pathway.
CC Phosphorylates beta-casein, histone 1 and myelin basic protein in
CC vitro.
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- COFACTOR: Magnesium.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated (By
CC similarity).
CC -1- PTM: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol
CC under basal conditions and dephosphorylated when membrane-
CC associated (By similarity).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: D49785; BAA08621.1; -
DR HSSP: P12931; IFMK.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR004040; STY_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00069; Pkinase.1.
DR ProDom: PD000001; Euk_pkinase.1.
DR SMART: SM00221; STYK; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation; Magnesium; Membrane.
FT DOMAIN 158 399
FT NP_BIND 164 172 ATP (BY SIMILARITY).
FT BINDING 185 185
FT ACT_SITE 269 269 BY SIMILARITY.
FT DOMAIN 56 62
FT DOMAIN 668 671
FT DOMAIN 698 701
FT DOMAIN 753 758
SQ SEQUENCE 888 AA; 96307 MW; 52AD964406BAE149 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 888;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRA 14
| | | | |
Db 98 PESRA 102

RESULT 182
KINH_NEUCR

ID	KINH_NEUCR	STANDARD;	PRT;	928 AA.
AC	P48467;			
DT	01-FEB-1996 (Rel. 33, Created)			
DT	15-JUL-1999 (Rel. 38, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	kinasin heavy chain.			
GN	KIN.			
OS	Neurospora crassa.			
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;			
OC	Sordariales; Sordariaceae; Neurospora.			
OX	NCBI_TaxID=5141;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.			
RC	STRAIN=74A;			
RX	MEDLINE=96157894; PubMed=8589459;			
RA	Steinberg G., Schliwa M.;			
RT	"The Neurospora organelle motor: a distant relative of conventional			
RT	kinasin with unconventional properties."			

RL Mol. Microbiol. 12:535-545(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-AC327;
 RX MEDLINE-96036197; PubMed-7581999;
 RA Rashid M.H., Mori M., Sekiguchi J.;
 RT "Glucosaminidase of Bacillus subtilis: cloning, regulation, primary
 structure and biochemical characterization.";
 RL Microbiology 141:2391-2404(1995).
 [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-168;
 RX MEDLINE-96044033; PubMed-9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azavedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borriss R., Boursier L., Brans A., Braun M., Brigelli S.C., Bron S.,
 RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel N.M.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Entian K.D., Erington J., Fabret C., Ferrari E., Foulger D.,
 RA Filtz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Gilm S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
 RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaert-Blanchard M., Klein C.,
 RA Kobayashi Y., Koeter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kunita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 RA Medina N., Melledo R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,
 RA Patro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Pirescen E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rago T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Saito T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Seliguchi J., Sekowaka A., Seror S.J., Serror P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takematu K.,
 RA Takeuchi M., Tanakoshi A., Tanaka T., Terpestra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassartotl A.,
 RA Viati A., Wambut R., Wedler E., Wedler H., Weltzenegger T.,
 RA Yoshida P., Yipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 subtilis.";
 RL Nature 390:249-256(1997).
 [2]
 RL -1- FUNCTION: CELL WALL HYDROLASE NOT INVOLVED IN CELL AUTOLYSIS. IT
 CC HYDROLYZES THE BETA-1,4 GLYCAN BOND BETWEEN THE N-
 CC ACETYLGLUCOSAMINYL AND THE N-ACETYLURAMONYL RESIDUES IN THE GLYCAN
 CC CHAIN.
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the di-N-acetylchitobiosyl
 CC unit in high-mannose glycopeptides and glycoproteins containing
 CC the -Man(GlcNAc)2)Asn-structure. One N-acetyl-D-glucosamine
 CC residue remains attached to the protein; the rest of the
 CC oligosaccharide is released intact.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO FAMILY 73 OF GLYCOSYL HYDROLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: 002562; AAA67857.1; -;
 DR EMBL: D45046; BA08089.1; -;
 DR EMBL: Z99122; CAB1595.1; -;
 DR Subtilist; BG10455; ltyd.
 DR InterPro: IPR002901; ltydase_4.
 DR InterPro: IPR003646; SH3_bac.
 DR Pfam: PF01832; Amidase_4; 1.
 DR SMART: SM00047; LY22; 1.

DR SMART: SM00287; SH3b; 1.
 KW Cell wall; Hydrolase; Signal; Repeat; Complete proteome.
 FT SIGNAL 1 27 POTENTIAL.
 FT CHAIN 28 880 BETA-N-ACETYLGLUCOSAMINIDASE.
 FT DOMAIN 72 75 POLY-THR.
 FT DOMAIN 337 340 POLY-LYS.
 FT DOMAIN 568 571 POLY-ALA.
 FT REPEAT 439 473 1.
 FT REPEAT 479 513
 SQ SEQUENCE 880 AA; 95553 MW; 2A912M478FCFC1D1 CRC64;
 Query Match 33.3%; Score 5; DB 1; Length 880;
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 7 RSTPE 11
 Db 643 RSTPE 647
 RESULT 180
 M3KC_MOUSE STANDARD: PRT; 888 AA.
 AC Q60700; P70286;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Mitogen-activated protein kinase kinase 12 (EC 2.7.1.37)
 DE (leucine-zipper protein kinase) (ZPK) (dual leucine zipper bearing
 DE kinase) (DLK).
 GN MAP3K12 OR ZPK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CD-1; TISSUE-Brain;
 RX MEDLINE-95074107; PubMed-7983011;
 RA Holzman L.B., Merritt S.E., Fan G.;
 RT "Identification, molecular cloning, and characterization of dual
 RT leucine zipper bearing kinase. A novel serine/threonine protein kinase
 RT that defines a second subfamily of mixed lineage kinases.";
 RL J. Biol. Chem. 269:30808-30817(1994).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ICR X Swiss Webster; TISSUE-Brain;
 RX MEDLINE-96365388; PubMed-8769565;
 RA Blouin R., Beaudoin J., Bergeron P., Nadeau A., Grondin G.;
 RT "Cell-specific expression of the ZPK gene in adult mouse tissues.";
 RL DNA Cell Biol. 15:631-642(1996).
 [3]
 RP PHOSPHORYLATION, AND MUTAGENESIS OF LYS-185 AND GLU-192.
 RX MEDLINE-96279269; PubMed-8663324.
 RA Mata M., Merritt S.E., Fan G., Yu G.G., Holzman L.B.;
 RT "Characterization of dual leucine zipper-bearing kinase, a mixed
 RT lineage kinase present in synaptic terminals whose phosphorylation
 RT state is regulated by membrane depolarization via calcineurin.";
 RL J. Biol. Chem. 271:16888-16896(1996).
 CC -1- FUNCTION: May be an activator of the JNK/SAPK pathway.
 CC Phosphorylates beta-casein, histone 1 and myelin basic protein in
 CC vitro.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein - ADP + a phosphoprotein.
 CC -1- COFACTOR: Magnesium.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated.
 CC -1- TISSUE SPECIFICITY: Expressed in brain, kidney, lung, heart,
 CC testis, gastrointestinal tract, stomach, liver and pancreas.
 CC Within the nervous system, predominantly expressed in neurons and
 CC enriched in synaptic terminals.
 CC -1- PTM: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol
 CC under basal conditions and dephosphorylated when membrane-
 CC associated.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AJ278263; CAC16404.2; ALT_INIT.
 DR EMBL: AJ276505; CAC27351.1; ALT_INIT.
 DR MGD: MGI:1930265; Oshp2.
 DR InterPro: IPR000648; Oxysterol_BP.
 DR InterPro: IPR001849; PH.
 DR Pfam: PF00169; PH; 1.
 DR Pfam: PF01237; Oxysterol_BP; 1.
 DR SMART: SM00233; PH; 1.
 DR PROSITE: PS01013; OSBP; 1.
 DR PROSITE: PS50003; PH_DOMAIN; 1.
 KW Lipid transport; Transport.
 FT DOMAIN 126 243 PH.
 FT CONFLICT 37 44 ENELGPIIT -> MSLVPSIQ (IN REF. 1;
 FT CAC27351).
 SQ SEQUENCE 874 AA; 98921 MW; FBC41FA8E219F5E3 CRC64;
 Query Match 33.3%; Score 5; DB 1; Length 874;
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 8 STEPS 12
 Db 779 STEPS 783

RESULT 178
 ORP5_HUMAN STANDARD; PRT; 879 AA.
 ID ORP5_HUMAN STANDARD; PRT; 879 AA.
 AC Q9H0X9; Q9H2B0; G9P124;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Oxysterol binding protein-related protein 5 (OSBP-related protein 5)
 DE (ORP-5)
 GN OSBP5 OR ORP5 OR KIAA1554.
 GN Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Annals A.M., Apostolopoulos J., Sparrow R.L.;
 RT "Isolation and characterization of human oxysterol-binding protein-
 RT related protein-5 (ORP-5).";
 RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21601154; PubMed=11735225;
 RA Jaworski C.J., Moreira E., Li A., Lee R., Rodriguez I.R.;
 RT "A family of 12 human genes containing oxysterol-binding domains.";
 RL Genomics 78:185-196(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Uterus;
 RX MEDLINE=21154917; PubMed=11230166;
 RA Wiemann S., Well B., Wellenreuther R., Gassenhuber J., Glassl S.,
 RA Ansgore W., Boecker M., Blocker H., Bauersachs S., Blum H.,
 RA Lauber J., Duesterhoeft A., Beyer R., Koehrer K., Strack N.,
 RA Mewes H.-W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
 RA Wambutt R., Korn B., Klein M., Poustka A.;
 RT "Towards a catalog of human genes and proteins: sequencing and
 RT analysis of 500 novel complete protein coding human cDNAs.";
 RL Genome Res. 11:422-435(2001).
 RN [4]
 RP SEQUENCE OF 1-61 FROM N.A.

RX MEDLINE=21376257; PubMed=11483621;
 RA Iehio M., Iaitinen S., Chinetti G., Johansson M., Ehnholm C.,
 RA Staels B., Ikonen E., Oikarinen V.M.;
 RT "The OSBP-related protein family in humans.";
 RL J. Lipid Res. 42:1203-1213(2001).
 RN [5]
 RP SEQUENCE OF 15-879 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=20277482; PubMed=10819331;
 RA Nagase T., Kikuno R., Ishikawa K.-I., Hirose M., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XVII.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 7:143-150(2000).
 CC -1- SIMILARITY: BELONGS TO THE OSBP FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AF410430; AAK98617.1; -
 DR EMBL: AF392453; AAL40666.1; -
 DR EMBL: AL136918; CAB6852.1; -
 DR EMBL: AF31964; AAG53417.1; -
 DR EMBL: AB040967; BAA96058.1; -
 DR Genew: HGNC:16392; OSBP5.
 DR MIM: 606733; -
 DR InterPro: IPR000648; Oxysterol_BP.
 DR InterPro: IPR001849; PH.
 DR Pfam: PF00169; PH; 1.
 DR Pfam: PF01237; Oxysterol_BP; 1.
 DR SMART: SM00233; PH; 1.
 DR PROSITE: PS01013; OSBP; 1.
 DR PROSITE: PS50003; PH_DOMAIN; 1.
 KW Lipid transport; Transport.
 FT DOMAIN 126 243 PH.
 FT CONFLICT 15 29 PPS5TPQKVDPRKLT -> EEGGHEKRRRRRAS (IN
 FT REF. 5).
 SQ SEQUENCE 879 AA; 98616 MW; 7EF06544347CC60A CRC64;
 Query Match 33.3%; Score 5; DB 1; Length 879;
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 8 STEPS 12
 Db 780 STEPS 784

RESULT 179
 LYTD_BACSU STANDARD; PRT; 880 AA.
 ID LYTD_BACSU STANDARD; PRT; 880 AA.
 AC P39848;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Beta-N-acetylglycosaminidase precursor (EC 3.2.1.96).
 GN LYTD OR CWG6.
 GN Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_Taxid=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=95020588; PubMed=7934877;
 RA Margot P., Maueel C., Karamata D.;
 RT "The gene of the N-acetylglycosaminidase, a Bacillus subtilis 168
 RT cell wall hydrolase not involved in vegetative cell autolysis.";

CC deformability by stabilizing spectrin-actin interaction. Binds
CC with a high affinity to glycophorin and with lower affinity to
CC band III protein.
CC -1- ALTERNATIVE PRODUCTS: A number of isoforms are produced by
CC alternative splicing.
CC -1- MASS SPECTROMETRY: MW:17199.3; METHOD-Electrospray; RANGE=709-858.
CC -1- SIMILARITY: CONTAINS 1 BAND 4.1-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L00919; AAA37123.1; -;
DR EMBL: L00919; AAA37122.1; -;
DR MGD: MGI:95401; EPD4.1.
DR InterPro: IPR000299; Band_4.1.
DR Pfam: PF00373; Band_41; 1.
DR PRINTS: PR00935; BAND41.
DR SMART: SM00295; BA1; 1.
DR PROSITE: PS00660; BAND_41_1; 1.
DR PROSITE: PS00661; BAND_41_2; 1.
DR PROSITE: PS50057; BAND_41_3; 1.
DR Structural protein; Alternative splicing; Cytoskeleton;
KW Phosphorylation.
FT DOMAIN 208 422 BAND 4.1-LIKE.
FT DOMAIN 489 608 HYDROPHILIC.
FT DOMAIN 609 707 SPECTRIN-ACTIN-BINDING.
FT DOMAIN 710 858 CARBOXYL-TERMINAL (CTD).
SQ SEQUENCE 858 AA; 95990 MW; 5F2FE077946134E CRC64;

Query Match 33.3%; Score 5; DB 1; Length 858;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRA 14
Db 190 PESRA 194
|||||
Db 190 PESRA 194

RESULT 176
M3KC_HUMAN STANDARD; PRT; 859 AA.
ID M3KC_HUMAN
AC 012852;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mitogen-activated protein kinase kinase 12 (EC 2.7.1.37)
DE (leucine-zipper protein kinase) (ZPK).
GN MAP3K12 OR ZPK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
ON NCBI_TaxID=9606;
RX NCBI_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=Testis; Carcinoma;
RX MEDLINE=94311945; PubMed=8037767;
RA Reddy U.R.; Pleasure D.;
RT "Cloning of a novel putative protein kinase having a leucine zipper
RT domain from human brain."
RL Blochm. Biophys. Res. Commun. 202:613-620(1994).
CC -1- FUNCTION: May be an activator of the JNK/SAPK pathway.
CC phosphorylates beta-casein, histone 1 and myelin basic protein in
CC vitro.
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- COFACTOR: Magnesium.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated (By
CC similarity).
CC -1- TISSUE SPECIFICITY: Highly expressed in brain and kidney.

CC -1- PTM: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol
CC under basal conditions and dephosphorylated when membrane-
CC associated (By similarity).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U07358; AAA67343.1; -;
DR HSSP: P12931; IPMK.
DR Genew: HGNC:6851; MAP3K12.
DR MIM: 600447; -;
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR004040; STY_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00221; STYKC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation; Magnesium; Membrane.
FT DOMAIN 125 366 PROTEIN KINASE.
FT NP_BIND 131 139 ATP (BY SIMILARITY).
FT BINDING 152 152 ATP (BY SIMILARITY).
FT ACT_SITE 236 236 BY SIMILARITY.
FT DOMAIN 665 668 POLY-PRO.
FT DOMAIN 720 725 POLY-GLU.
SQ SEQUENCE 859 AA; 93188 MW; 0E5209792C5C6F05 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 859;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRA 14
Db 65 PESRA 69
|||||
Db 65 PESRA 69

RESULT 177
ORP5_MOUSE STANDARD; PRT; 874 AA.
ID ORP5_MOUSE
AC Q9ER64; Q99NP5;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Oxyesterol binding protein-related protein 5 (OSBP-related protein 5)
DE (ORP-5) (Oxyesterol-binding protein homologue 1).
GN OSBP5 OR OSBP2 OR OBPH1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RX NCBI_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RX MEDLINE=20519229; PubMed=11063728;
RA Engemann S.; Stroedecke M.; Paulsen M.; Franck O.; Reinhardt R.;
RA Lane N.; Reik W.; Walter J.;
RT "Sequence and functional comparison in the Beckwith-Wiedemann region:
RT implications for a novel imprinting centre and extended imprinting."
RL Hum. Mol. Genet. 9:2691-2706(2000).
CC -1- SIMILARITY: BELONGS TO THE OSBP FAMILY.
CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC		TRANSDUCTION.
CC	-1-	SUBCELLULAR LOCATION: Integral membrane protein.
CC	-1-	TISSUE SPECIFICITY: KIDNEY AND INTESTINE.
CC	-1-	PWM: PHOSPHORYLATED (POSSIBLE).
CC	-1-	SMILIARITY: BELONGS TO THE NA(+/H(+)) EXCHANGER FAMILY.
CC	-1-	CATION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC DOMAINS IN THE NA(+)/H(+) EXCHANGERS VARY AMONG AUTHORS.
CC		-----
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC		-----
DR	EMBL; X59935;	CAA42258.1; -
DR	EMBL; X61504;	CAA43721.1; -
DR	EMBL; X56536;	CAA39881.1; -
DR	PIR; S13926;	S13926.
DR	PIR; S16328;	S16328.
DR	InterPro; IPR000676;	NAH_Exchange.
DR	InterPro; IPR004709;	NAH_exchange3.
DR	Pfam; PF009099;	Na_H_Exchange; 1.
DR	PRINTS; PR01084;	NAHEXCHNGR.
DR	TIGRFAMS; TIGR00840;	b_cpal; 1.
KW	Transmembrane; Glycoprotein; Transport; Antipor; Sodium transport;	
KW	Multigene family; Phosphorylation.	
FT	DOMAIN	1 15 CYTOPLASMIC (POTENTIAL).
FT	TRANSSEM	16 35 M1 (POTENTIAL).
FT	DOMAIN	36 107 EXTRACELLULAR (POTENTIAL).
FT	TRANSSEM	108 127 M2 (POTENTIAL).
FT	DOMAIN	128 129 CYTOPLASMIC (POTENTIAL).
FT	TRANSSEM	130 149 M3 (POTENTIAL).
FT	DOMAIN	150 154 EXTRACELLULAR (POTENTIAL).
FT	TRANSSEM	155 174 M4 (POTENTIAL).
FT	DOMAIN	175 191 CYTOPLASMIC (POTENTIAL).
FT	TRANSSEM	192 211 M5 (POTENTIAL).
FT	DOMAIN	212 227 EXTRACELLULAR (POTENTIAL).
FT	TRANSSEM	228 247 M5A (POTENTIAL).
FT	DOMAIN	248 256 CYTOPLASMIC (POTENTIAL).
FT	TRANSSEM	257 276 M5B (POTENTIAL).
FT	DOMAIN	277 294 EXTRACELLULAR (POTENTIAL).
FT	TRANSSEM	295 315 M6 (POTENTIAL).
FT	DOMAIN	316 338 CYTOPLASMIC (POTENTIAL).
FT	TRANSSEM	339 358 M7 (POTENTIAL).
FT	DOMAIN	359 386 EXTRACELLULAR (POTENTIAL).
FT	TRANSSEM	387 406 M8 (POTENTIAL).
FT	DOMAIN	407 410 CYTOPLASMIC (POTENTIAL).
FT	TRANSSEM	411 430 M9 (POTENTIAL).
FT	DOMAIN	431 480 EXTRACELLULAR (POTENTIAL).
FT	TRANSSEM	481 500 M10 (POTENTIAL).
FT	DOMAIN	501 816 CYTOPLASMIC (POTENTIAL).
FT	CARBOHYD	75 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	370 370 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CONFLICT	242 242 V -> A (IN REF. 2).
FT	CONFLICT	569 569 K -> E (IN REF. 2).
SQ	SEQUENCE	816 AA; 90717 MW; 336738D267F7436 CRC64;
	Query Match	33.3%; Score 5; DB 1; Length 816;
	Best Local Similarity	100.0%; Pred. No. 3;le+02;
	Matches 5; Conservative	0; Mismatches 0; Indels 0; Gaps 0.
OY	9 TPESR 13	
DB	68 TPESR 72	
RESULT 172	BGAL_ASPOF	STANDARD; PRT; 832 AA.
ID	P45582;	
DT	01-NOV-1995 (Rel. 32, Created)	

DT	01-NOV-1995 (Rel. 32, Last sequence update)
DT	16-OCT-2001 (Rel. 40, Last annotation update)
DE	Beta-galactosidase precursor (EC 3.2.1.23) (Lactase).
OS	Asparagus officinalis (Garden asparagus).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC	Spermatophytas; Magnoliophyta; Liliopsida; Asparagales; Asparagaceae;
CC	Asparagus
OX	NCBI_TaxID=4686;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=cv. Limbras 10; TISSUE=Spear;
RX	MEDLINE=95303968; PubMed=7784512.
RA	King G.A., Davies K.M.;
RT	"Cloning of a harvest-induced beta-galactosidase from tips of
RT	harvested asparagus spears."
RL	Plant Physiol. 108:419-420(1995).
CC	-1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-
CC	galactose residues in beta-D-galactosides.
CC	-1- SIMILARITY: BELONGS TO FAMILY 35 OF GLYCOSYL HYDROLASES.
CC	-1- SIMILARITY: CONTAINS 1 SUEL-TYPE LECTIN DOMAIN.
CC	-----
CC	This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL: X77319; CAA54525.1; -.
DR	InterPro: IPR001944; GH_35.
DR	InterPro: IPR000922; Gal_lectin.
DR	Pfam: PF01301; Glyco_hydro_35; 1.
DR	Pfam: PF02140; Gal_lectin; 1.
DR	PRINTS: PRO0742; GLHYDRLASE35.
DR	ProDom: PD005612; Gal_lectin; 1.
DR	PROSITE: PS01182; GLYCOSYL_HYDROL_F35; 1.
DR	PROSITE: PS50228; SUEL_LECTIN_1.
RW	Hydrolase; Glycosidase; Signal.
FT	CHAIN 1 25
FT	SIGNAL 1 25
FT	POTENTIAL.
FT	BETA-GALACTOSIDASE.
FT	SUEL-TYPE LECTIN.
FT	PROTON DONOR (POTENTIAL).
FT	ACT_SITE 741 832
FT	ACT_SITE 183 183
FT	ACT_SITE 252 252
FT	NCLEOPHILE (POTENTIAL).
SO	SEQUENCE 832 AA; 92213 MW; 94ABDC61BCA4164AE CRC64;
Query Match	33.3%; Score 5; DB 1; Length 832;
Best Local Similarity	100.0%; Pred. No. 3; le+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	7 RSTPE 11
Db	52 RSTPE 56
RESULT 173	
BGAL_LYCES	
ID_BGAL_LYCES	STANDARD; PRT; 835 AA.
AC	P48980;
DT	01-FEB-1996 (Rel. 33, Created)
DT	01-FEB-1996 (Rel. 33, Last sequence update)
DT	15-JUN-2002 (Rel. 41, Last annotation update)
DE	Beta-galactosidase precursor (EC 3.2.1.23) (Lactase) (Acid beta-
DE	galactosidase) (Exo-(1-->4)-beta-D-galactanase).
OS	Lycopersicon esculentum (tomato)
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophytas; Magnoliophyta; eudicotyledons; core eudicots;
CC	Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX	NCBI_TaxID=4081;
RN	[1]
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC	STRAIN=cv. Alisa Craig; TISSUE=pericarp.
RX	MEDLINE=95357407; PubMed=7630937;

RA Garden O.A., Musk P., Worthington-White D.A., Dewey M.J., Rich I.N.;
 RT "Silent polymorphisms within the coding region of human
 RT sodium/hydrogen exchanger isoform-1 cDNA in peripheral blood
 RT mononuclear cells of leukemia patients: A comparison with healthy
 RT controls.";
 RL Cancer Genet. Cytogenet. 120:37-43(2000).
 CC -1- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED
 CC BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL
 CC CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD
 CC SODIUM ION CHEMICAL GRADIENT. PLAYS AN IMPORTANT ROLE IN SIGNAL
 CC TRANSDUCTION.
 CC -1- SUBUNIT: Interacts with tescalcin.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: KIDNEY AND INTESTINE.
 CC -1- PTM: PHOSPHORYLATED (POSSIBLE).
 CC -1- MISCELLANEOUS: INHIBITED BY AMILORIDE AND 5-AMINO-SUBSTITUTED
 CC DERIVATIVES AND ACTIVATED IN A COOPERATIVE FASHION BY
 CC INTRACELLULAR H⁺. FULLY ACTIVE AT ACIDIC PH, THE ANTI-PORTER IS
 CC VIRTUALLY TURNED OFF AT NEUTRAL PH. IN QUISCENT CELLS UPON GROWTH
 CC FACTOR STIMULATION, THE APPARENT AFFINITY FOR INTRNAL H⁺ IS
 CC INCREASED, RESULTING IN A PERSISTENT RISE IN CYTOPLASMIC PH.
 CC -1- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.
 CC -1- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC
 CC DOMAINS IN THE NA(+)/H(+) EXCHANGERS VARY AMONG AUTHORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M81768; AAB59460.1; ALT_SEQ.
 DR EMBL: S68616; AAC60606.1; -
 DR EMBL: AF141350; AAF21350.1; -
 DR EMBL: AF141351; AAF21351.1; -
 DR EMBL: AF141352; AAF21352.1; -
 DR EMBL: AF141353; AAF21353.1; -
 DR EMBL: AF141354; AAF21354.1; -
 DR EMBL: AF141355; AAF21355.1; -
 DR EMBL: AF141356; AAF21356.1; -
 DR EMBL: AF141357; AAF21357.1; -
 DR EMBL: AF141358; AAF21358.1; -
 DR EMBL: AF141359; AAF21359.1; -
 DR EMBL: AF146430; AAF25592.1; -
 DR EMBL: AF146431; AAF25593.1; -
 DR EMBL: AF146432; AAF25594.1; -
 DR EMBL: AF146433; AAF25595.1; -
 DR EMBL: AF146434; AAF25596.1; -
 DR EMBL: AF146435; AAF25597.1; -
 DR EMBL: AF146436; AAF25598.1; -
 DR EMBL: AF146437; AAF25599.1; -
 DR EMBL: AF146438; AAF25600.1; -
 DR EMBL: AF146439; AAF25601.1; -
 DR PIR: A31311; A31311.
 DR GENE: HGNC:11071; SLC9A1.
 DR MIM: 107310; -
 DR InterPro: IPR000676; Nah_Exchange.
 DR InterPro: IPR004709; Nah_exchanger3.
 DR Pfam: PF00999; Nah_H_Exchange; 1.
 DR PRINTS: PRO1084; NAHEXCHNGR.
 DR TIGR: TIGR00840; b_cpai; 1.
 KW Transmembrane; Glycoprotein; Transport; Antiport; Sodium transport;
 KW Multigene family; Phosphorylation.
 FT DOMAIN 1 15 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 16 35 M1 (POTENTIAL).
 FT DOMAIN 36 107 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 108 127 M2 (POTENTIAL).
 FT DOMAIN 128 129 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 130 149 M3 (POTENTIAL).
 FT DOMAIN 150 154 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 155 174 M4 (POTENTIAL).

FT DOMAIN 175 191 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 192 211 M5 (POTENTIAL).
 FT DOMAIN 212 227 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 228 247 M5A (POTENTIAL).
 FT DOMAIN 248 256 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 257 276 M5B (POTENTIAL).
 FT DOMAIN 277 294 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 295 315 M6 (POTENTIAL).
 FT DOMAIN 316 338 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 339 358 M7 (POTENTIAL).
 FT DOMAIN 359 386 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 387 406 M8 (POTENTIAL).
 FT DOMAIN 407 410 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 411 430 M9 (POTENTIAL).
 FT DOMAIN 431 480 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 481 500 M10 (POTENTIAL).
 FT DOMAIN 501 815 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 75 75 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CARBOHYD 370 370 N-LINKED (GLCNAc. . .) (POTENTIAL).
 SQ SEQUENCE 815 AA; 90763 MW; 02EC748C79DF6526 CRC64;
 Query Match Score 5; DB 1; Length 815;
 Best Local Similarity 100.0%; Pred. No. 3; 1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 9 TPESR 13
 DB 68 TPESR 72
 ID NAHL_RABBIT STANDARD; PRT; 816 AA.
 AC P23791;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sodium/hydrogen exchanger 1 (Na(+)/H(+) exchanger 1) (NHE-1).
 GN SLC9A1 OR NHE1.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-New Zealand white; TISSUE=Ileal villus;
 RX MEDLINE=91293066; PubMed=1712287;
 RA Tse C.-M., Ma A.I., Yang V.W., Watson A.J.M., Levine S.,
 RA Montrose M.H., Potter J., Sargent C., Pouyssegur J., Donowitz M.;
 RT "Molecular cloning and expression of a cDNA encoding the rabbit ileal
 RT villus cell basolateral membrane Na⁺/H⁺ exchanger.";
 RL EMO J. 10:1957-1967(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=92096447; PubMed=1661611;
 RA Hildebrandt F., Pizzonia J.H., Rellly R.F., Reboucas N.A.,
 RA Sargent C., Pouyssegur J., Slayman C.W., Aronson P.S., Jgarashi P.;
 RT "Cloning, sequence, and tissue distribution of a rabbit renal Na⁺/H⁺
 RT exchanger transcript.";
 RL Biochim. Biophys. Acta 1129:105-108(1991).
 RN [3]
 RP SEQUENCE OF 472-816 FROM N.A.
 RC STRAIN-New Zealand white; TISSUE=Heart muscle;
 RX MEDLINE=91138752; PubMed=1704856;
 RA Fliegel L., Sargent C., Pouyssegur J., Barr A.;
 RT "Identification of the protein and cDNA of the cardiac Na⁺/H⁺
 RT exchanger.";
 RL FEBS Lett. 279:25-29(1991).
 CC -1- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED
 CC BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL
 CC CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD
 CC SODIUM ION CHEMICAL GRADIENT. PLAYS AN IMPORTANT ROLE IN SIGNAL

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001644; AAD18733.1; -
DR EMBL: AE002176; AAF8035.1; -
DR EMBL: AP002547; BAA98801.1; -
DR TIGR: CP0154; -
DR InterPro: IPR005146; B3_4.
DR InterPro: IPR005147; B5.
DR InterPro: IPR005121; Fdx-AnticB.
DR InterPro: IPR004532; Pher-bact.
DR InterPro: IPR002547; tRNA_bind.
DR Pfam: PF01588; tRNA_bind; 1.
DR Pfam: PF03483; B3_4; 1.
DR Pfam: PF03484; B5; 1.
DR TIGRFAMs: TIGR00472; phet_bact; 2.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT CONFLICT 171 171 G -> A (IN REF. 3).
FT CONFLICT 577 577 C -> Y (IN REF. 1).
SQ SEQUENCE 792 AA; 87732 MW; 852C86BF0F07486F CRC64;

Query Match 33.3%; Score 5; DB 1; Length 792;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 8 STRES 12
|||||
Db 285 STRES 289

RESULT 169
ID SECA_MYCPN STANDARD: PRT; 808 AA.
AC P75559;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Preprotein translocase seca subunit.
GN SECA OR MPN210 OR MP621.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelfeich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -1- FUNCTION: INVOLVED IN PROTEIN EXPORT. INTERACTS WITH THE SECY/SECE
CC SUBUNTS. SECA HAS A CENTRAL ROLE IN COUPLING THE HYDROLYSIS OF
CC ATP TO THE TRANSFER OF PRE-SECRETORY PERIPLASMIC AND OUTER
CC MEMBRANE PROTEINS ACROSS THE MEMBRANE (BY SIMILARITY).
CC -1- SUBUNIT: PART OF THE PROKARYOTIC PROTEIN TRANSLOCATION APPARATUS
CC WHICH COMPRISE SECA, SECB, SECD, SECE, SECF, SECG AND SECY
CC (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC SIDE OF MEMBRANE
CC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SECA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified, and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000059; AAB96269.1; -
DR InterPro: IPR001650; Helicase_C.
DR InterPro: IPR000185; SecA.
DR Pfam: PF00271; helicase_C; 1.
DR Pfam: PF01043; SecA_protein; 1.
DR PRINTS: PR00906; SECA.
DR TIGRFAMs: TIGR00963; seca; 1.
DR PROSITE: PS01312; SECA; 1.
KW Protein transport; ATP-binding; Membrane; Translocation; Transport;
KW Complete proteome.
FT NP_BIND 102
SQ SEQUENCE 808 AA; 91810 MW; E768F0B8915F892F CRC64;

Query Match 33.3%; Score 5; DB 1; Length 808;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 10 PESRA 14
|||||
Db 255 PESRA 259

RESULT 170
ID NAHL_HUMAN STANDARD: PRT; 815 AA.
AC P19634;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sodium/hydrogen exchanger 1 (Na(+)/H(+) exchanger 1) (NHE-1) (Na+/H+
DE antiporter, amiloride-sensitive) (APNH).
GN SLC9A1 OR NHE1 OR APNH1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Kidney;
RX MEDLINE=89106219; PubMed=2536298;
RA Sargent C., Franchi A., Pouyssegur J.;
RT "Molecular cloning, primary structure, and expression of the human
RT growth factor-activatable Na+/H+ antiporter.";
RL Cell 56:271-280(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Kidney;
RX MEDLINE=90140739; PubMed=2154036;
RA Sargent C., Connillon L., Franchi A., Pouyssegur J.;
RT "Growth factors induce phosphorylation of the Na+/H+ antiporter,
RT glycoprotein of 110 kD.";
RL Science 247:723-726(1990).
RN [3]
RP REVISIONS.
RX MEDLINE=91293066; PubMed=1712287;
RA Tse C.-M., Ma A.T., Yang V.W., Watson A.J.M., Levine S.,
RA Montrose M.H., Potter J., Sargent C., Pouyssegur J., Donowitz M.;
RT "Molecular cloning and expression of a cDNA encoding the rabbit ileal
RT villus cell basolateral membrane Na+/H+ exchanger.";
RL EMBO J. 10:1957-1967(1991).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=94111706; PubMed=8283968;
RA Fliegel L., Dyck J.R., Wang H., Fong C., Haworth R.S.;
RT "Cloning and analysis of the human myocardial Na+/H+ exchanger.";
RL Mol. Cell. Biochem. 125:137-143(1993).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=20375279; PubMed=10913675;

DR InterPro: IPR000201; DNAPol_viral_N.
DR InterPro: IPR000477; RVase.
DR Pfam: PF00078; rvt; 1.
DR Pfam: PF00242; DNA_pol_viral_N; 1.
DR Pfam: PF00336; DNA_pol_viral_C; 1.
DR ProDom: PD000814; DNAPol_viral_C; 1.
KW Transferrase; RNA-directed DNA polymerase; DNA-directed DNA polymerase;
KW Hydrolyase; Nuclease; Endonuclease; DNA replication; DNA-binding.
SQ SEQUENCE 788 AA; 89378 MW; F482FBS78D5B5B1B CRC64;

Query Match 33.3%; Score 5; DB 1; Length 788;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRA 14
DB 355 PESRA 359

RESULT 167
CADI_HUMAN STANDARD; PRT; 790 AA.
ID CADI_HUMAN
AC 013634;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cadherin-18 precursor (Cadherin-14).
GN CDH18 OR CDH14.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA MEDLINE=97184182; PubMed=9030594;
RA Shibata T., Shimoyama Y., Gotoh M., Hirohashi S.;
RT "Identification of human cadherin-14, a novel neurally specific type
II cadherin, by protein interaction cloning.";
RL J. Biol. Chem. 272:5236-5240(1997).
CC -!- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
CC THEIR PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE
CC SORTING OF HETEROGENEOUS CELL TYPES.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U59325; AAB02933.1; -
CC HSSP: P15116; INCJ.
CC Genew: HGNC:1757; CDH18.
CC MIM: 603019;
DR InterPro: IPR002126; Cadherin.
DR InterPro: IPR000233; Cadherin_C_term.
DR Pfam: PF00028; cadherin; 5.
DR Pfam: PF01049; Cadherin_C_term; 1.
DR PRINTS: PRO0205; CADHERIN.
DR SMART: SM00112; CA; 5.
DR PROSITE: PS00232; CADHERIN_1; 3.
DR PROSITE: PS50268; CADHERIN_2; 5.
KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
KW Signal.
FT SIGNAL 1 24
FT PROPEP 25 53 POTENTIAL.
FT CHAIN 54 790 CADHERIN-18.
FT DOMAIN 54 608 EXTRACELLULAR (POTENTIAL).
CC

FT TRANSMEM 609 636 POTENTIAL.
FT DOMAIN 637 790 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 54 159 CADHERIN 1.
FT DOMAIN 160 268 CADHERIN 2.
FT DOMAIN 269 383 CADHERIN 3.
FT DOMAIN 384 486 CADHERIN 4.
FT DOMAIN 487 608 CADHERIN 5.
FT CARBOHYD 36 36 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 255 255 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 455 455 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 536 536 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 790 AA; 88072 MW; 5C7BDEB229B6EDCA CRC64;

Query Match 33.3%; Score 5; DB 1; Length 790;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGP 5
DB 359 SHLGP 363

RESULT 168
SYFB_CHLPN STANDARD; PRT; 792 AA.
ID SYFB_CHLPN
AC 0927W0; Q9K2D4; Q9JSD3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Phenylalanyl-tRNA synthetase beta chain (EC 6.1.1.20) (Phenylalanyl--
DE tRNA ligase beta chain) (PheRS).
GN PHER OR CPN0594 OR CP0154.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RA MEDLINE=9920606; PubMed=10192388;
RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RA MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Ullrich T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwin M., Nelson W., Deboy R., Koloney J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RA MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hiraoka H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kohara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CW1029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
CC -!- CATALYTIC ACTIVITY: ATP + L-phenylalanine + tRNA(Phe) -> AMP +
CC diphosphate + L-phenylalanyl-tRNA(Phe).
CC -!- SUBUNIT: Tetramer of two alpha and two beta chains (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE PHENYLALANYL-TRNA SYNTHETASE BETA
CC CHAIN FAMILY. SUBFAMILY 1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

DB      542 STPES 546

RESULT 164
K6PF_ASPNG      STANDARD:      PRT:      783 AA.
AC      P78985;
DR      15-JUL-1998 (Rel. 36, Created)
DT      15-JUL-1998 (Rel. 36, Last sequence update)
DE      16-OCT-2001 (Rel. 40, Last annotation update)
DE      6-phosphofructokinase (EC 2.7.1.11) (Phosphofructokinase)
DE      (Phosphohexokinase).
GN      PFKA.
OS      Aspergillus niger.
OC      Eukaryota: Fungi: Ascomycota: Pezizomycotina: Eurotiomycetes:
OC      Eurotiiales: Trichocommataceae; mitosporic Trichocommataceae; Aspergillius.
OX      NCBI_TaxID=5061;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=CBS 120.49 / N400;
RA      de Graaff L.H., Everse S.J., van den Broeck H.C., Bussink R.,
RA      Visser J.;
RL      Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC      -1- CATALYTIC ACTIVITY: ATP + D-fructose 6-phosphate = ADP + D-
CC      fructose 1,6-bisphosphate.
CC      -1- PATHWAY: KEY CONTROL STEP OF GLYCOLYSIS.
CC      -1- SIMILARITY: BELONGS TO THE PHOSPHOFRUCTOKINASE FAMILY. TWO DOMAIN
CC      SUBFAMILY.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
CC      EMBL: 279690; CAB01923.1; -.
CC      HSSP: P00512; 3PRK.
CC      InterPro: IPR000023; Pfrfructokinase.
CC      DR      Pfam: PF00365; PFK: 2
CC      PRINTS: PR00476; PFRFRCTKINASE.
CC      PRODOM: PD000707; Pfrfructokinase: 2.
CC      DR      PROSITE: PS00433; PHOSPHOFRUCTOKINASE: 2.
CC      KW      Transferase; Kinase; Glycolysis; Repeat.
CC      SEQUENCE 783 AA; 85759 MW; 7BE3F6B53F06304 CRC64;

Query Match      33.3%; Score 5; DB 1; Length 783;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      11 ESRAA 15
DB      666 ESRAA 670

RESULT 165
DFOL_HPBD      STANDARD:      PRT:      787 AA.
AC      P30028;
DR      01-APR-1993 (Rel. 25, Created)
DT      01-APR-1993 (Rel. 25, Last sequence update)
DE      16-OCT-2001 (Rel. 40, Last annotation update)
DE      P protein [Includes: DNA-directed DNA polymerase (EC 2.7.7.7); RNA-
DE      directed DNA polymerase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN      P.
OS      Duck hepatitis B virus (strain China) (DHBV).
OC      Viruses: Retroid viruses; Hepadnaviridae; Avihepadnavirus.
OX      NCBI_TaxID=31510;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      MEDLINE=91045091; PubMed=2235506;
RA      Tong S., Matles F., Teubner K., Blum H.E.;

```

```

RT      "Complete nucleotide sequence of a Chinese duck hepatitis B virus.";
RL      Nucleic Acids Res. 18:6139-6139(1990).
CC      -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate
CC      + (DNA)(N).
CC      -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC      phosphomonoester.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
CC      EMBL: M21953; AAA5745.1; -.
CC      PIR: S12841; JDIVLW2.
CC      DR      InterPro: IPR001462; DNAPol_viral_C.
CC      DR      InterPro: IPR000201; DNAPol_viral_N.
CC      DR      InterPro: IPR000477; RVase.
CC      DR      Pfam: PF00028; rvt; 1.
CC      DR      Pfam: PF00242; DNA_pol_viral_N; 1.
CC      DR      Pfam: PF00336; DNA_pol_viral_C; 1.
CC      DR      PRODOM: PD000814; DNAPol_viral_C; 1.
CC      KW      Transferase; RNA-directed DNA polymerase; DNA-directed DNA polymerase;
CC      Hydrolyase; Nuclease; Endonuclease; DNA replication; DNA-binding.
CC      SEQUENCE 787 AA; 89214 MW; 1A99D7A656665180 CRC64;

Query Match      33.3%; Score 5; DB 1; Length 787;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      10 PESRA 14
DB      355 PESRA 359

RESULT 166
DFOL_HPBD      STANDARD:      PRT:      788 AA.
AC      P17192;
DR      01-AUG-1990 (Rel. 15, Created)
DT      01-AUG-1990 (Rel. 15, Last sequence update)
DE      16-OCT-2001 (Rel. 40, Last annotation update)
DE      P protein [Includes: DNA-directed DNA polymerase (EC 2.7.7.7); RNA-
DE      directed DNA polymerase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN      P.
OS      Duck hepatitis B virus (brown Shanghai duck isolate S5) (DHBV).
OC      Viruses: Retroid viruses; Hepadnaviridae; Avihepadnavirus.
OX      NCBI_TaxID=10439;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      MEDLINE=90085807; PubMed=2596031;
RA      Uchida M., Esumi M., Shikata T.;
RA      "Molecular cloning and sequence analysis of duck hepatitis B virus
RA      genomes of a new variant isolated from Shanghai ducks.";
RL      Virology 173:600-606(1989).
CC      -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate
CC      + (DNA)(N).
CC      -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC      phosphomonoester.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
CC      EMBL: M32990; AAA5754.1; -.
CC      PIR: A33746; JDIVLBD.
CC      DR      InterPro: IPR001462; DNAPol_viral_C.

```



```

FT1 REPEAT 359 364 9.
FT1 REPEAT 365 370 10.
FT1 REPEAT 371 376 11.
FT1 REPEAT 377 382 12.
FT1 REPEAT 383 388 13.
FT1 REPEAT 389 394 14.
FT1 REPEAT 395 400 15.
FT1 REPEAT 401 406 16.
FT1 REPEAT 407 412 17.
FT1 REPEAT 413 418 18.
FT1 REPEAT 419 424 19.
FT1 REPEAT 425 429 20 (IMPERFECT).
FT1 REPEAT 430 435 21.
FT1 REPEAT 436 441 22.
FT1 DOMAIN 466 666 MAGE.
FT1 CONFLICT 178 178 O -> R (IN REF. 3).
FT1 CONFLICT 425 427 CIL -> WIP (IN REF. 1).
FT1 CONFLICT 693 699 LETECL -> IGDEAVS (IN REF. 1).
FT1 CONFLICT 724 725 WS -> GP (IN REF. 3).
SQ SEQUENCE 775 AA; 85861 MW; 2E554B7B365512E1 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 775;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STRES 12
|||||
Db 539 STRES 543

RESULT 163
MGDI_HUMAN STANDARD; PRT; 778 AA.
AC 09Y5Y3; 09UP36; 09HBT4; 09H352;
ID 16-OCT-2001 (Rel. 40, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Melanoma-associated antigen D1 (MAGE-D1 antigen) (Neurotrophin
DE receptor-interacting MAGE homolog) (PRO2292).
GN MAGE1 OR NRAGE.
OS Homo sapiens (Human).
OC Bkariyoda; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCB1_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=bone marrow;
RX MEDLINE=99339980; PubMed=10409427;
RA Foid M., Zhou J., Chen G.L., Hall J.M., Vescio R.A., Berenson J.R.;
RT "Identification of a new, unorthodox member of the MAGE gene family.";
RL Genomics 59:161-167(1999).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20439481; PubMed=10985348;
RA Salehi A.H., Roux P.P., Kudu C.J., Zeindler C., Bhakar A.,
RA Tanais L.-L., Verdi J.M., Barker P.A.;
RT "NRAGE, a novel MAGE protein, interacts with the p75 neurotrophin
RT receptor and facilitates nerve growth factor dependent apoptosis.";
RL Neuron 27:279-288(2000).
[3]
RP SEQUENCE OF 304-778 FROM N.A.
RC TISSUE=Testis;
RA Blum H., Bauersachs S., Mewes H.-W., Gassenhuber J., Wiemann S.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
[4]
RP SEQUENCE OF 396-778 FROM N.A.
RC TISSUE=Fetal liver;
RX MEDLINE=21177478; PubMed=11280991;
RA Zhang C.G., Xing G.C., Wei H.D., Yu Y.T., He F.C.;
RT "A new melanoma antigen encoding gene subfamily in human chromosome
RT X.";
RL I Chuan Hsueh Pao 28:197-203(2001).
[5]

```

```

RP IDENTIFICATION OF THE TRANSLATIONAL INITIATION CODON.
RX MEDLINE=20541720; PubMed=11087672;
RA Kudu C.J., Goldhawk D.G., Barker P.A., Verdi J.M.;
RT "Identification of the translational initiation codon in human
RT MAGE1.";
RL Genomics 70:150-152(2000).
CC -1- FUNCTION: Involved in the apoptotic response after nerve growth
CC factor (NGF) binding in neuronal cells. Binds p75NTR and
CC antagonizes its association with TrkA. Inhibits cell cycle
CC progression, and facilitates p75NTR-mediated apoptosis. May act as
CC a regulator of the function of DLX family members.
CC -1- SUBUNIT: Interacts with the p75 neurotrophin receptor.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic. Expression shifts from the
CC cytoplasm to the plasma membrane upon stimulation with NGF (By
CC similarity).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN BONE MARROW STROMAL CELLS FROM
CC BOTH MULTIPLE MYELOMA PATIENTS AND HEALTHY DONORS. SEEMS TO BE
CC UBQUITOUSLY EXPRESSED.
CC -1- SIMILARITY: CONTAINS 1 MAGE DOMAIN.
CC -1- CAUTION: REF.1 differs from that shown due to several frameshifts
CC that resulted in a N-terminally truncated protein.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF124440; AAC31421.1; ALT_FRAME.
CC DR EMBL; AF217963; AAG09704.1; -.
CC DR EMBL; AL133628; CAB63752.1; -.
CC DR EMBL; AF132205; AAC35551.1; ALT_INT.
CC DR Genew; HGNC:6813; MAGE1.
CC DR MIM; 300224; -.
CC DR InterPro; IPR002190; MAGE.
CC DR Pfam; PF01454; MAGE; 3.
CC DR PROSITE; PS50838; MAGE; 1.
CC KW Antigen; Multigene family; Repeat.
CC FT DOMAIN 296 444 22 X 6 AA TANDEM REPEATS OF W-[PQ]-X-P-X-
CC FT REPEAT 296 301 1.
CC FT REPEAT 302 307 2.
CC FT REPEAT 308 313 3.
CC FT REPEAT 332 337 4.
CC FT REPEAT 338 343 5.
CC FT REPEAT 344 349 6.
CC FT REPEAT 350 355 7.
CC FT REPEAT 356 361 8.
CC FT REPEAT 362 367 9.
CC FT REPEAT 368 373 10.
CC FT REPEAT 374 379 11.
CC FT REPEAT 380 385 12.
CC FT REPEAT 386 391 13.
CC FT REPEAT 392 397 14.
CC FT REPEAT 398 403 15.
CC FT REPEAT 404 409 16.
CC FT REPEAT 410 415 17.
CC FT REPEAT 416 421 18.
CC FT REPEAT 422 427 19.
CC FT REPEAT 428 432 20 (IMPERFECT).
CC FT REPEAT 433 438 21.
CC FT REPEAT 439 444 22.
CC FT DOMAIN 471 669 MAGE.
SQ SEQUENCE 778 AA; 86150 MW; 0F8BEC7155326FCC CRC64;

Query Match 33.3%; Score 5; DB 1; Length 778;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STRES 12
|||||

```



```

CC cytoplasm to the plasma membrane upon stimulation with NGF (By
CC similarity).
CC -1- TISSUE SPECIFICITY: Ubiquitously expressed in many adult tissues,
CC except for the spleen. Expressed in osteoblastic and
CC chondrogenic cell lines and also during embryonic development.
CC -1- SIMILARITY: CONSTRAINTS 1 MAGE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB029448; BAA87959.1; -
DR EMBL: AF319975; AAK01203.1; -
DR EMBL: AK017275; BAB30666.1; -
DR EMBL: AK013231; BAB28729.1; ALT_INIT.
DR MGD: MGI:1930187; Maged1.
DR InterPro: IPR002190; MAGE.
DR Pfam: PF01454; MAGE; 1.
DR PROSITE: PS50838; MAGE; 1.
KW Antigen; Multigene family; Repeat.
FT DOMAIN 292 441 22 X 6 AA TANDEM REPEATS OF W-[PQ]-X-P-X-
FT REPEAT 292 297 1.
FT REPEAT 298 303 2.
FT REPEAT 304 309 3.
FT REPEAT 329 334 4.
FT REPEAT 335 340 5.
FT REPEAT 341 346 6.
FT REPEAT 347 352 7.
FT REPEAT 353 358 8.
FT REPEAT 359 364 9.
FT REPEAT 365 370 10.
FT REPEAT 371 376 11.
FT REPEAT 377 382 12.
FT REPEAT 383 388 13.
FT REPEAT 389 394 14.
FT REPEAT 395 400 15.
FT REPEAT 401 406 16.
FT REPEAT 407 412 17.
FT REPEAT 413 418 18.
FT REPEAT 419 424 19.
FT REPEAT 425 429 20 (IMPERFECT).
FT REPEAT 430 435 21.
FT REPEAT 436 441 22.
FT DOMAIN 468 666 MAGE.
FT CONFLICT 357 362 MISSING (IN REF. 2).
SQ SEQUENCE 775 AA; 85669 MW; 224B82470816835A CRC64;
Query Match 33.3%; Score 5; DB 1; Length 775;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 8 STPES 12
DB 539 STPES 543

```

```

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CC NCBI_Taxid:10116;
CC [1]
CC SEQUENCE FROM N.A.
CC STRAIN-Sprague-Dawley; TISSUE-Neural crest;
CC MEDLINE=20439481; PubMed=10985348;
CC Salehi A.H., Roux P.P., Kubu C.J., Zeindler C., Bhakar A.,
CC Tannis L.-L., Verdi J.M., Barker P.A.;
CC "NRAGE", a novel MAGE protein, interacts with the p75 neurotrophin
CC receptor and facilitates nerve growth factor dependent apoptosis.";
CC Neuron 27:279-288(2000).
CC [2]
CC SEQUENCE OF 206-775 FROM N.A.
CC TISSUE-Testis;
CC MEDLINE=20466167; PubMed=11014239;
CC Henny B., Reiter E., Cornet A., Bruyninx M., Daukandt M., Housa P.,
CC N'Guyen V.-H., Closset J., Hennen G.;
CC "A novel messenger ribonucleic acid homologous to human MAGE-D is
CC strongly expressed in rat Sertoli cells and weakly in Leydig cells
CC and is regulated by follitropin, lutropin, and prolactin.";
CC Endocrinology 141:3821-3831(2000).
CC [3]
CC SEQUENCE OF 178-775 FROM N.A.
CC Zhang C., He F.;
CC "Rattus norvegicus mRNA for SNERG-1 protein, partial CDS.";
CC Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: Involved in the apoptotic response after nerve growth
CC factor (NGF) binding in neuronal cells. Binds p75NTR and
CC antagonizes its association with TrkA, inhibits cell cycle
CC progression, and facilitates p75NTR-mediated apoptosis. May act as
CC a regulator of the function of Dlx family members.
CC -1- SUBUNIT: INTERACTS WITH DLX5, DLX7 AND MSX2 AND FORMS
CC HOMOMULTIMERS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic. Expression shifts from the
CC cytoplasm to the plasma membrane upon stimulation with NGF.
CC -1- TISSUE SPECIFICITY: Ubiquitous and in the seminiferous tubules
CC expressed in sertoli cells but not in germ cells. Expression
CC decreases in all tissues with increased age and is detectable only
CC in brain cortex and lung.
CC -1- DEVELOPMENTAL STAGE: Expressed at low levels throughout the embryo
CC and is enriched in the developing brain and spinal cord.
CC -1- INDUCTION: Follitropin decreased expression while lutropin and
CC prolactin stimulated expression.
CC -1- SIMILARITY: CONSTRAINTS 1 MAGE DOMAIN.
CC -1- CAUTION: REF 2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 726.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF217964; AAG09705.1; -
DR EMBL: AJ133038; CAB65381.1; ALT_FRAME.
DR EMBL: AF274043; AAF75283.1; -
DR InterPro: IPR002190; MAGE.
DR Pfam: PF01454; MAGE; 3.
DR PROSITE: PS50838; MAGE; 1.
KW Antigen; Multigene family; Repeat.
FT DOMAIN 293 441 22 X 6 AA TANDEM REPEATS OF W-[PQ]-X-P-X-
FT REPEAT 293 298 1.
FT REPEAT 299 304 2.
FT REPEAT 305 310 3.
FT REPEAT 329 334 4.
FT REPEAT 335 340 5.
FT REPEAT 341 346 6.
FT REPEAT 347 352 7.
FT REPEAT 353 358 8.

```

```

KW Hydrolase; Glycosidase; Periplasmic; Signal; Complete proteome.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 765 PERIPLASMIC BETA-GLUCOSIDASE.
FT ACT_SITE 287 287 BY SIMILARITY.
SQ SEQUENCE 765 AA; 83460 MW; 0D89B0AB42B8F8F3 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 765;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGP 5
Dp 364 SHLGP 368

RESULT 160
BGLX_SALTY STANDARD; PRT; 765 AA.
ID BGLX_SALTY
AC 056078;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Periplasmic beta-glucosidase precursor (EC 3.2.1.21) (Gentiobiose)
DE (Cellulobiose) (beta-D-glucoside glucosylhydrolase) (T-cell inhibitor).
GN BGLX OR STM2166.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RA Matsui K.;
RL Submitted (JUG-1996) to the EMBL/GenBank/DBJ databases.
RT [2]
RN SEQUENCE FROM N.A.
RC STRAIN=LT2 / SCSCL12 / ATCC 700720;
RA McDellane-21534948; PubMed-11677609;
RA McLelland M., Sanderson K.E., Spiehl J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grevail N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2."
RL Nature 413:852-856(2001).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-
CC glucose residues with release of beta-D-glucose.
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF GLYCOSYL HYDROLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: D86507; BAA13102.1; -
DR EMBL: AE008796; AAL21070.1; -
DR Stuygen; SG10604; bglX.
DR InterPro: IPR002772; GH_3C.
DR InterPro: IPR001764; GH_3N.
DR Pfam: PF00933; Glyco_hydro_3; 1.
DR Pfam: PF01915; Glyco_hydro_3_C; 1.
DR PROSITE: PS00775; GLYCOSYL_HYDROL_F3; 1.
KW Hydrolase; Glycosidase; Periplasmic; Signal; Complete proteome.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 765 PERIPLASMIC BETA-GLUCOSIDASE.
FT ACT_SITE 287 287 BY SIMILARITY.
FT CONFLICT 630 630 R -> L (IN REF. 1).
FT CONFLICT 634 635 EA -> KP (IN REF. 1).

```

```

FT CONFLICT 656 656 V -> F (IN REF. 1).
FT CONFLICT 674 674 E -> K (IN REF. 1).
FT CONFLICT 717 717 E -> K (IN REF. 1).
SQ SEQUENCE 765 AA; 83392 MW; C7C767EBAC96786 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 765;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGP 5
Dp 364 SHLGP 368

RESULT 161
MGDL_MOUSE STANDARD; PRT; 775 AA.
ID MGDL_MOUSE
AC 09QYH6; 099PB5; 09CXY1;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Melanoma-associated antigen D1 (MAGE-D1 antigen) (Neurotrophin
DE receptor-interacting MAGE homolog) (Dlx1-1).
GN MAGE-D1 OR NRAGE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RC MEDLINE=21265065; PubMed=11084035;
RA Masuda Y., Sasaki A., Shibuya H., Ueno N., Ikeda K., Matanabe K.;
RT "Dlx1-1, a novel protein that binds Dlx5 and regulates its
RT transcriptional function."
RL J. Biol. Chem. 276:5331-5338(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Anguier P.H., Chomez P.M., De Backer O.R., Bertrand M.J.M.;
RT "Ten new murine members of the MAGE gene family."
RT Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Akawa T., Hara A., Fukunishi Y., Komoto H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamahata I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gasterland T., Glassl C., King B., Kochli H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirni L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boileau D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofman M., Hume D.A., Kamaya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momhaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
CC -1- FUNCTION: Involved in the apoptotic response after nerve growth
CC factor (NGF) binding in neuronal cells. Binds p75NTR and
CC antagonizes its association with TrkA, inhibits cell cycle
CC progression, and facilitates p75NTR-mediated apoptosis. May act as
CC a regulator of the function of Dlx family members (By similarity).
CC -1- SUBUNIT: INTERACTS WITH DLX5, DLX7 AND MSX2 AND FORMS
CC HOMODIMERS.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic. Expression shifts from the

```

RA Premont R.T., Ciaing A., Vitale N., Perry S.J., Lefkowitz R.J.;
 RT "The G12 family of ADP-ribosylation factor GTPase-activating proteins.
 RT Functional diversity of G12 through alternative splicing.";
 RL J. Biol. Chem. 275:22373-22380(2000).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE-96127530; PubMed-8590280;
 RA Nagase T., Seki N., Tanaka A., Ishikawa K.-I., Nomura N.;
 RT "Prediction of the coding sequences of unidentified human genes. IV.
 RT The coding sequences of 40 new genes (K1AA0121-K1AA0160) deduced by
 RT analysis of cDNA clones from human cell line KG-1.";
 RL DNA Res. 2:167-174(1995).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE-SKIN;
 RA Strausberg R.;
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP INTERACTION WITH PAXILLIN.
 RX MEDLINE-21149839; PubMed-11251077;
 RA Masaki Y., Hashimoto S., Okawa K., Tsudouchi A., Nakamura K., Yagi R.,
 RA Yano H., Kondo A., Iwanatsu A., Mizoguchi A., Sabe H.;
 RT "An ADP-ribosylation factor GTPase-activating protein Glt2-
 RT short/K1AA0148 is involved in subcellular localization of paxillin and
 RT actin cytoskeletal organization.";
 RL Mol. Biol. Cell 12:645-662(2001).
 CC -1- FUNCTION: GTPase-activating protein for the ADP ribosylation
 CC factor family.
 CC -1- SUBUNIT: Interacts with G protein-coupled receptor kinases.
 CC Associates with paxillin. Also interacts with PIX exchange
 CC factors.
 CC -1- ALTERNATIVE PRODUCTS: At least 9 isoforms; 1 (shown here), 2/G12-
 CC short, 3/C-, 4/BC-, 5/E-, 6/CD-, 7/DE-, 8/BE-, 9/AE-; are produced
 CC by alternative splicing. Many more seems to be produced by
 CC skipping of internal exons.
 CC -1- SIMILARITY: CONTAINS 1 ARF-GAP DOMAIN.
 CC -1- SIMILARITY: CONTAINS 3 ANK REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: A124491; AAD28047.1; -
 DR EMBL: D63482; BAA09769.1; -
 DR EMBL: BC001379; AAH01379.1; -
 DR Genew: HGNC:4273; G12.
 DR InterPro: IPR002110; ANK.
 DR InterPro: IPR001164; hRIP_like.
 DR Pfam: PF000023; ank, 3.
 DR Pfam: PF01412; ArfGAP, 1.
 DR PRINTS: PR00405; REVINTRACTNG.
 DR SMART: SM00248; ANK, 3.
 DR SMART: SM00105; ArfGAP, 1.
 DR PROSITE: PS50088; ANK_REPEAT, 1.
 DR PROSITE: PS50297; ANK_REPEAT_REGION, 1.
 DR PROSITE: PS50115; ARFGAP, 1.
 KW GTPase activation; Repeat; ANK repeat; Zinc-finger;
 KW Alternative splicing.
 FT DOMAIN 1 124
 FT ZN_FING 11 34 ARF-GAP.
 FT REPEAT 132 161 C4-TYPE.
 FT REPEAT 166 195 ANK 1.
 FT REPEAT 199 228 ANK 2.
 FT REPEAT 466 471 ANK 3.
 FT VARSPPLIC 472 471 OTIOSE -> LGKDN (IN ISOFORM 2).
 FT VARSPPLIC 472 759 MISSING (IN ISOFORM 2).
 FT VARSPPLIC 334 414 MISSING (IN ISOFORM 9).
 FT VARSPPLIC 415 449 MISSING (IN ISOFORM 4 AND ISOFORM 8).
 FT VARSPPLIC 450 464 MISSING (IN ISOFORM 3, ISOFORM 4 AND
 FT ISOFORM 6).
 FT

FT VARSPPLIC 465 547 MISSING (IN ISOFORM 6 AND ISOFORM 7).
 FT VARSPPLIC 548 577 MISSING (IN ISOFORM 5, ISOFORM 7,
 FT CONFLICT 285 285 MISSING (IN ISOFORM 9).
 FT SEQUENCE 759 AA; 84542 MW; 07FE266DB2F3258 CRC64;
 V -> M (IN REF. 3).
 Query Match 33.3%; Score 5; DB 1; Length 759;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 8 STRES 12
 DB 586 STRES 590
 RESULT 159
 BGX_ECOLI STANDARD; PRT; 765 AA.
 ID BGX_ECOLI
 AC P33363;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Periplasmic beta-glucosidase precursor (EC 3.2.1.21) (Gentloblase)
 DE (Cellubiose) (Beta-D-glucoside gluconhydrolase).
 GN BGX OR B2132.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / W3110;
 RA Yang M., Luo S., Goddard A., Reilly D., Henzel W., Bass S.;
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / BHB2600;
 RA Richterich P., Lakey N., Gryan G., Jaehn L., Mintz L., Robison K.,
 RA Church G.M.;
 RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE-97426617; PubMed-9278503;
 RA Blatter F.R., Plunkett G. II, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-
 CC glucose residues with release of beta-D-glucose.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF GLYCOSYL HYDROLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: U15049; AAB38487.1; -
 DR EMBL: U00007; AAB60495.1; ALT_INIT.
 DR EMBL: AE000302; AAC75193.1; -
 DR EcoGene: EGI2013; BglX.
 DR InterPro: IPR002772; GH_3C.
 DR InterPro: IPR001764; GH_3N.
 DR Pfam: PF00933; Glyco_hydro_3, 1.
 DR Pfam: PF01915; Glyco_hydro_3_C, 1.
 DR PRINTS: PR00133; GLHYDRLASE3.
 DR PROSITE: PS00775; GLYCOSYL_HYDROL_F3, 1.
 DR

DT 15-JUN-2002 (Rel. 41, last annotation update)
 DE Cytosolic phospholipase A2 (cPLA2) [includes: Phospholipase A2
 DE (EC 3.1.1.4) (phosphatidylcholine 2-acylhydrolase); Lysophospholipase
 DE (EC 3.1.1.5)].
 GN PLA2G4A OR PLA2G4 OR CPLA2.
 OS Brachydanio rerio (Zedratishn) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 NCBI_TaxID=7955;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=9429545; PubMed=8027085;
 RA Nalefski E.A., Sultzman L.A., Martin D.M., Kriz R.W., Towler P.S.,
 RA Knopf J.L., Clark J.D.;
 RT "Delineation of two functionally distinct domains of cytosolic
 RT phospholipase A2, a regulatory Ca(2+)-dependent lipid-binding domain
 RT and a Ca(2+)-independent catalytic domain.";
 RL J. Biol. Chem. 269:18239-18249(1994).
 CC -1- FUNCTION: SELECTIVELY HYDROLYZES ARACHIDONYL PHOSPHOLIPIDS IN THE
 CC SN-2 POSITION RELEASING ARACHIDONIC ACID. TOGETHER WITH ITS
 CC LIPOPHOSPHOLIPID ACTIVITY, IT IS IMPLICATED IN THE INITIATION OF
 CC THE INFLAMMATORY RESPONSE.
 CC -1- CATALYTIC ACTIVITY: Phosphatidylcholine + H(2)O = 1-
 CC acylglycerophosphocholine + a fatty acid anion.
 CC -1- CATALYTIC ACTIVITY: 2-lysophosphatidylcholine + H(2)O =
 CC glycerophosphocholine + a fatty acid anion.
 CC -1- ENZYME REGULATION: STIMULATED BY AGONISTS SUCH AS ATP, EGF,
 CC THROMBIN AND BRADYKININ AS WELL AS BY CYTOSOLIC CA2+.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC. TRANSLOCATES TO MEMBRANE
 CC VESICLES IN A CALCIUM-DEPENDENT FASHION (BY SIMILARITY).
 CC -1- DOMAIN: THE N-TERMINAL C2 DOMAIN, BY ITS ASSOCIATION WITH LIPID
 CC MEMBRANES, MEDIATES THE REGULATION OF cPLA2 BY PRESENTING THE
 CC ACTIVE SITE TO ITS SUBSTRATE IN RESPONSE TO ELEVATIONS OF
 CC CYTOSOLIC CA2+ (BY SIMILARITY).
 CC -1- PTM: ACTIVATED BY PHOSPHORYLATION ON A SERINE RESIDUE (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U10330; AAA53229.1; -
 CC HSSP: P47712; IBCI.
 CC ZFIN: ZDB-GENE-990415-45; cpla2.
 CC InterPro: IPR000008; C2.
 CC InterPro: IPR002642; PLAC.
 CC Pfam: PF00168; C2; 1.
 CC Pfam: PF01735; PLA2_B; 1.
 CC SMART: SM00239; C2; 1.
 CC SMART: SM00022; PLAC; 1.
 CC PROSITE: PS50004; C2_DOMAIN_2; 1.
 KW Hydrolyase; Lipid degradation; Calcium; Phosphorylation.
 FT DOMAIN 1 172 PHOSPHOLIPID BINDING (PROBABLE).
 FT MOD_RES 498 498 C2 DOMAIN.
 FT PHOSPHORYLATION (BY MAPK) (BY
 FT SIMILARITY).
 SQ SEQUENCE 741 AA; 83809 MW; 34896B1A8364A9D4 CRC64;

RESULT 157
 YK09_YEAST
 ID YK09_YEAST STANDARD; PRT; 751 AA.
 AC P36124;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, last annotation update)
 DE Hypothetical 85.5 kDa protein in SAP190-SP014 intergenic region.
 GN YK0929C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 NCBI_TaxID=4932;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA Urrestarazu L.A., Jauniaux J.-C.;
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO YEAST YJL105W AND S.POMBE SPAC22E12.11C.
 CC -1- SIMILARITY: CONTAINS 1 PHD-TYPE ZINC FINGER.
 CC -1- SIMILARITY: CONTAINS 1 SET DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Z28254; CAAB2101.1; -
 CC PIR: S38101; S38101.
 CC SCD: S0001737; YK0929C.
 CC InterPro: IPR001214; SET.
 CC InterPro: IPR001965; Znf_PHD.
 CC Pfam: PF00628; PHD; 1.
 CC Pfam: PF00856; SET; 1.
 CC SMART: SM00249; PHD; 1.
 CC SMART: SM00317; SET; 1.
 CC PROSITE: PS50280; SET; 1.
 CC PROSITE: PS01359; ZF_PHD_1; 1.
 CC PROSITE: PS50016; ZF_PHD_2; 1.
 KW Hypothetical protein; Zinc-finger.
 FT ZN_FING 117 166 PHD-TYPE.
 FT DOMAIN 334 460 SET.
 SQ SEQUENCE 751 AA; 85479 MW; 934621768C362308 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 751;
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STEPS 12
 DB 718 STEPS 722

RESULT 158
 GIT2_HUMAN
 ID GIT2_HUMAN STANDARD; PRT; 759 AA.
 AC Q14161; Q9Y5V2; Q9BV91;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 16-OCT-2001 (Rel. 40, last sequence update)
 DT 15-JUN-2002 (Rel. 41, last annotation update)
 DE ARF GTPase-activating protein GIT2 (G protein-coupled receptor kinase-
 DE interactor 2).
 GN GIT2 OR KIAA0148.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RX [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 TO 9), AND CHARACTERIZATION.
 RX MEDLINE=20357364; PubMed=10896954;

[7]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 25-220.
 RA Wang J.-H., Stehle T., Peplinsky R.B., Liu J.-H., Karpusas M.,
 RA Osborn L.:
 RT "Structure of a functional fragment of VCAM-1 refined at 1.9-A
 RT resolution."
 RL Acta Crystallogr. D 52:369-379(1996).
 CC -1- FUNCTION: IMPORTANT IN CELL-CELL RECOGNITION. APPEARS TO FUNCTION
 CC IN LEUKOCYTE-ENDOTHELIAL CELL ADHESION. INTERACTS WITH THE BETA-1
 CC INTEGRIN VIA4 ON LEUKOCYTES, AND MEDIATES BOTH ADHESION AND SIGNAL
 CC TRANSDUCTION. THE VCAM1/VIA4 INTERACTION MAY PLAY A
 CC PATHOPHYSIOLOGIC ROLE BOTH IN IMMUNE RESPONSES AND IN LEUKOCYTE
 CC EMIGRATION TO SITES OF INFLAMMATION.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS: A LONG FORM (SHOWN
 CC HERE) AND A SHORT FORM: ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: EXPRESSED ON INFLAMED VASCULAR ENDOTHELIUM, AS
 CC WELL AS ON MACROPHAGE-LIKE AND DENDRITIC CELL TYPES IN BOTH NORMAL
 CC AND INFLAMED TISSUE.
 CC -1- INDUCTION: BY CYTOKINES (E.G. IL-1, TNF-ALPHA).
 CC -1- PTM: SIALOGLYCOPROTEIN.
 CC -1- DISEASE: MAY PLAY AN IMPORTANT ROLE IN THE GENESIS OF
 CC ARTEROSCLEROSIS AND RHEMATOID ARTHRITIS.
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
 CC -1- SIMILARITY: CONTAINS 7 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -1- DATABASE: NAME=PROV; NOTE=CD guide CD106 entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd106.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X53051; CA37218.1; -
 DR EMBL: M30257; AA51917.1; ALT-TERM.
 DR EMBL: M73255; AA61270.1; -
 DR EMBL: M60335; AA61269.1; -
 DR PIR: A33758; A33758.
 DR PIR: A39755; A39755.
 DR PIR: A41288; A41288.
 DR PIR: B41288; B41288.
 DR PIR: S11476; S11476.
 DR PIR: IVCA; 15-SEP-95.
 DR PDB: 1VCA; 20-JUN-96.
 DR Gene: HGNC:12663; VCAM1.
 DR MIM: 192225; -
 DR InterPro: IPR003987; ICAM_VCAM-1.
 DR InterPro: IPR003006; IG_MHC.
 DR InterPro: IPR003598; IG_C2.
 DR InterPro: IPR003600; IG_Like.
 DR InterPro: IPR003989; VCAM-1.
 DR Pfam: PF00047; 1g. 6.
 DR PRINTS: PR01472; ICAMVCAM1.
 DR PRINTS: PR01474; VCAM1.
 DR SMART: SM00410; IG_Like; 2.
 DR SMART: SM00408; IGC2; 3.
 KW Immunoglobulin domain; Glycoprotein; Cell adhesion; Transmembrane;
 KW Repeat; Signal; Alternative splicing; 3D-structure.
 FT SIGNAL 1 24
 FT CHAIN 25 739
 FT DOMAIN 25 698
 FT TRANSMEM 699 720
 FT DOMAIN 721 739
 FT DOMAIN 38 106
 FT DOMAIN 129 202
 FT DOMAIN 237 287
 FT DOMAIN 326 379
 FT DOMAIN 418 496
 FT DOMAIN 525 575
 FT DOMAIN 612 675

FT DISULFID 47 95
 FT DISULFID 52 99
 FT DISULFID 137 195
 FT CARBOHYD 273 273
 FT CARBOHYD 365 365
 FT CARBOHYD 417 417
 FT CARBOHYD 463 463
 FT CARBOHYD 531 531
 FT CARBOHYD 561 561
 FT VARSPLIC 310 402
 SO SEQUENCE 739 AA; 81276 MW; 050E2BD39AC2FF4 CRC64;
 Query Match 33.3%; Score 5; DB 1; Length 739;
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 9 TPESR 13
 DB 30 TPESR 34
 RESULT 155
 BSG2_DROME STANDARD; PRT; 741 AA.
 ID BSG2_DROME
 AC P11929;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 01-MAR-1992 (Rel. 21, Last annotation update)
 DE Blastoderm specific protein 25D.
 GN BSG25D.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 CC Muscophora; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=87174755; PubMed=3104878;
 RX Boyer P.D., Mahoney P.A., Lengyel J.A.;
 RT "Molecular characterization of bsg25D: a blastoderm-specific locus of
 RT Drosophila melanogaster."
 RL Nucleic Acids Res. 15:2309-2325(1987).
 CC -1- DEVELOPMENTAL STAGE: BLASTODERM SPECIFIC.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X04896; CA28582.1; -
 DR PIR: A26572; A26572.
 DR Flybase: FBgn0000228; Bsg25D.
 KW Developmental protein.
 SO SEQUENCE 741 AA; 84428 MW; 9ED8CE9F7EE2CD9F CRC64;
 Query Match 33.3%; Score 5; DB 1; Length 741;
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 9 TPESR 13
 DB 538 TPESR 542
 RESULT 156
 ID PA24_BRARE STANDARD; PRT; 741 AA.
 AC P50392;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)

```

CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L29451; AAA62324.1; -.
CC InterPro: IPR001944; GH_35.
CC Pfam: PF01301; Glyco_hydro_35; 1.
CC PRINTS: PR00742; GHYDRLASE35.
CC PROSITE: PS01182; GLYCOSYL_HYDROL_F35; 1.
CC Hydrolase; Glycosidase; Signal.
CC SIGNAL 1 23
CC CHAIN 24 731
CC ACT_SITE 182 182
CC ACT_SITE 251 251
CC ACT_SITE 731 AA; 80995 MW; FAB5D24A0D30BD4 CRC64;
SQ SEQUENCE

Query Match 33.3%; Score 5; DB 1; Length 731;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 STRPE 11
DB 51 STRPE 55

RESULT 153
NIBL_HUMAN STANDARD; PRT; 733 AA.
AC 09NT35; 09BUS2;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Niban-like protein (Meg-3).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RA Miyata T., Inagi R., Yasuda Y., Kurokawa K.;
RT "Homo sapiens meg-3 mRNA, complete cds."
RL Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.
RN [2]
RN SEQUENCE OF 38-733 FROM N.A.
RC TISSUE=Testis;
RA Ottenweider B., Obermaler B., Mewes H.-W., Gassenhuber J.,
RA Wiemann S.;
RT Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
RN [3]
RN SEQUENCE OF 278-733 FROM N.A.
RC TISSUE=Cervix;
RA Strausberg R.;
RL Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE NIBAN FAMILY.
CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF151783; AAK57556.1; -.
CC EMBL: AL137555; CAB70809.1; -.
CC EMBL: BC001979; AAH01979.1; ALT_INIT.
CC InterPro: IPR001849; PH.
CC Pfam: PF00169; PH; 1.
CC SMART: SM00233; PH; 1.
CC PROSITE: PS50003; PH_DOMAIN; 1.
CC DOMAIN 55 179
CC CONFLICT 305 305
CC CONFLICT 664 664
CC SEQUENCE 733 AA; 82683 MW; 38B1C24CF737E3DB CRC64;

```

```

Query Match 33.3%; Score 5; DB 1; Length 733;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STRPE 12
DB 592 STRPE 596

RESULT 154
NIBL_HUMAN STANDARD; PRT; 739 AA.
AC P19320;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Vascular cell adhesion protein 1 precursor (VCAM-1) (CD106 antigen)
DE (INCAM-100).
DE (INCAM-100).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RC TISSUE=umbilical vein;
RA MEDLINE-91016951; PubMed-1699207;
RA Polte T., Newman W., Gopal T.V.;
RT "Full length vascular cell adhesion molecule 1 (VCAM-1).";
RL Nucleic Acids Res. 18:5901-5901(1990).
RN [2]
RN SEQUENCE FROM N.A.
RA MEDLINE-90090619; PubMed-2688898;
RA Osborn L., Hession C., Tizard R., Vassallo C., Luhnowsky S.,
RA Chi-Rosso G., Lobb R.;
RT "Direct expression cloning of vascular cell adhesion molecule 1, a
RT cytokine-induced endothelial protein that binds to lymphocytes.";
RL Cell 59:1203-1211(1989).
RN [3]
RN SEQUENCE FROM N.A.
RA MEDLINE-91352090; PubMed-1715583;
RA Cybulsky M.I., Fries J.W.U., Williams A.J., Sultan P., Eddy R.,
RA Byers M., Shows T., Gimbrone M.A. Jr., Collins T.;
RT "Gene structure, chromosomal location, and basis for alternative mRNA
RT splicing of the human VCAM1 gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:7859-7863(1991).
RN [4]
RN SEQUENCE FROM N.A.
RA MEDLINE-91201302; PubMed-1707873;
RA Hession C., Tizard R., Vassallo C., Schiffer S.B., Goff D., Moy P.,
RA Chi-Rosso G., Luhnowsky S., Lobb R., Osborn L.;
RT "Cloning of an alternate form of vascular cell adhesion molecule-1
RT (VCAM1).";
RL J. Biol. Chem. 266:6682-6685(1991).
RN [5]
RN X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 25-226.
RA MEDLINE-95147978; PubMed-7531291;
RA Jones E.Y., Harlos K., Bottomley M.J., Robinson R.C., Driscoll P.C.,
RA Edwards R.M., Clements J.M., Dudgeon T.J., Stuart D.I.;
RT "Crystal structure of an integrin-binding fragment of vascular cell
RT adhesion molecule-1 at 1.8-A resolution.";
RL Nature 373:539-544(1995).
RN [6]
RN X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 25-220.
RA MEDLINE-95296382; PubMed-7539925;
RA Wang J.-H., Pepinsky R.B., Stehle T., Liu J.-H., Karpusas M.,
RA Browning B., Osborn L.;
RT "The crystal structure of an N-terminal two-domain fragment of
RT vascular cell adhesion molecule 1 (VCAM-1): a cyclic peptide based on
RT the domain 1 C-D loop can inhibit VCAM-1-alpha 4 integrin
RT interaction.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:5714-5718(1995).

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L06434; AAB24776.1; -.
DR EMBL: S68944; AAC60673.1; -.
DR PIR: S27043; S27043.
DR InterPro: IPR00175; Na/ntrn_symport.
DR Pfam: PF00209; SNF; 1.
DR PRINTS: PR00176; NANEUSMPORT.
DR PRODOM: PD000448; Na/ntrn_symport; 2.
DR PROSITE: PS00610; NA_NEUTROTAN_SYM_1; 1.
DR PROSITE: PS00754; NA_NEUTROTAN_SYM_2; 1.
DR PROSITE: PS50267; NA_NEUTROTAN_SYM_3; 1.
KW Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW Symport.
FT DOMAIN 1 68 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 69 89 1 (POTENTIAL).
FT TRANSMEM 97 116 2 (POTENTIAL).
FT TRANSMEM 141 161 3 (POTENTIAL).
FT DOMAIN 162 224 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 225 243 4 (POTENTIAL).
FT TRANSMEM 252 269 5 (POTENTIAL).
FT TRANSMEM 305 322 6 (POTENTIAL).
FT TRANSMEM 334 355 7 (POTENTIAL).
FT DOMAIN 356 451 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 452 471 8 (POTENTIAL).
FT TRANSMEM 495 513 9 (POTENTIAL).
FT TRANSMEM 529 549 10 (POTENTIAL).
FT TRANSMEM 570 591 11 (POTENTIAL).
FT TRANSMEM 619 641 12 (POTENTIAL).
FT DOMAIN 642 727 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 393 393 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 261 261 Y -> S (IN REF. 2).
FT CONFLICT 499 499 G -> S (IN REF. 2).
SO SEQUENCE 727 AA; 81055 MW; C676048C0A6DF7C CRC64;

Query Match 33.3%; Score 5; DB 1; Length 727;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STPES 12
Db 721 STPES 725

RESULT 151
BGAL_DIRACA STANDARD; PRT; 731 AA.
AC Q00662;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative beta-galactosidase precursor (EC 3.2.1.23) (Lactase)
DE (Srl2 protein).
GN CARSRL2.
OS Dianthus caryophyllus (Carnation) (Clove pink).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Caryophyllaceae; Dianthus.
OC NCBI_TaxID=3570;
RX MEDLINE=91329738; PubMed=1868223;
RA Raghoebar K.G., Lawton K.A., Goldsbrough P.B., Woodson W.R.;
RT "Characterization of an ethylene-regulated flower senescence-related

```

```

RT gene from carnation.";
RL Plant Mol. Biol. 17:61-71(1991).
CC -I CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-
CC galactose residues in beta-D-galactosides.
CC -I TISSUE SPECIFICITY: SENESCING FLOWER PETALS.
CC -I INDUCTION: BY ETHYLENE.
CC -I SIMILARITY: BELONGS TO FAMILY 35 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X57171; CAA40459.1; -.
DR PIR: S16595; S16595.
DR InterPro: IPR001944; GH_35.
DR Pfam: PF01301; Glyco_hydro_35; 1.
DR PRINTS: PR00742; GLHYDRASE35.
DR PROSITE: PS01182; GLYCOSYL_HYDROL_F35; 1.
KW Hydrolase; Glycosidase; Signal.
FT SIGNAL 1 29 POTENTIAL.
FT CHAIN 30 731 PUTATIVE BETA-GALACTOSIDASE.
FT ACT_SITE 187 187 PROTON DONOR (POTENTIAL).
FT ACT_SITE 257 257 NUCLEOPHILE (POTENTIAL).
SO SEQUENCE 731 AA; 82864 MW; 83FAB5A3779C051 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 731;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPE 11
Db 56 RSTPE 60

RESULT 152
BGAL_MALDO STANDARD; PRT; 731 AA.
AC P46981;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Beta-galactosidase precursor (EC 3.2.1.23) (Lactase) (Acid beta-
DE galactosidase) (Exo-(1-->4)-beta-D-galactanase).
OS Malus domestica (Apple) (Malus sylvestris).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Rosales; Rosaceae; Maloideae; Malus.
OC NCBI_TaxID=3750;
RX MEDLINE=95083752; PubMed=7991682;
RA Ross G.S., Megarzyn T., Macrae E.A., Redgwell R.J.;
RT "Apple beta-galactosidase. Activity against cell wall polysaccharides
RT and characterization of a related cDNA clone."
RL Plant Physiol. 106:521-528(1994).
CC -I FUNCTION: Involved in cell wall degradation. Degrades
CC polysaccharides containing beta-(1-->4)-linked galactans, acting
CC as an exo-(1-->4)-beta-D-galactanase.
CC -I CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-
CC galactose residues in beta-D-galactosides.
CC -I SIMILARITY: BELONGS TO FAMILY 35 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```



```

RN [2]
RP REVISION TO 654.
RA Requena J.M.;
RL Submitted (APR-2002) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: MOLECULAR CHAPERONE. HAS ATPASE ACTIVITY
CC (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 90 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X87770; CAD50506.1; -.
DR HSSP: P07900; IYER.
DR InterPro: IPR003594; ATPbind_ATPase.
DR InterPro: IPR001404; HSP90.
DR Pfam: PF00183; HSP90_1.
DR Pfam: PF02518; HATPase_C_1.
DR PRINTS: PR00775; HEATSHOCK90.
DR SMART: SM00387; HATPase_C_1.
DR PROSITE: PS00298; HSP90_1.
KW Chaperone; ATP-binding; Heat shock.
SQ SEQUENCE 701 AA; 80679 MW; A7FF41FF5264BF08 CRC64;

Query Match
Best Local Similarity 33.3%; Score 5; DB 1; Length 701;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STRIPS 12
DB 159 STRIPS 163

RESULT 149
GIT2_MOUSE
ID GIT2_MOUSE STANDARD; PRT; 708 AA.
AC Q9UJQ2;
DT 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE ARF GTPase-activating protein GIT2 (G protein-coupled receptor kinase-
DE interactor 2) (Tyrosine-phosphorylated protein Cat-2).
GN GIT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9351767; PubMed=10428811;
RA Bagrodia S., Bailey D., Lenard Z., Hart M., Guan J.L., Premont R.T.,
RA Taylor S.J., Cerione R.A.;
RT "A tyrosine-phosphorylated protein that binds to an important
RT regulatory region on the cool family of p21-activated kinase-binding
RT proteins."
RL J. Biol. Chem. 274:22393-22400(1999).
CC -1- FUNCTION: GTPase-activating protein for the ADP ribosylation
CC factor family (By similarity).
CC -1- SUBUNIT: Interacts with G protein-coupled receptor kinases.
CC Associates with paxillin. Also interacts with PIX exchange
CC factors (By similarity).
CC -1- PTM: Tyrosine-phosphorylated when co-expressed in cells with the
CC focal adhesion kinase Fak and Src.
CC -1- SIMILARITY: CONTAINS 1 ARF-GAP DOMAIN.
CC -1- SIMILARITY: CONTAINS 3 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF148693; AAF61633.1; -.
DR MGD: MGI:1347053; Gtl2.
DR InterPro: IPR002110; ANK.
DR InterPro: IPR001164; hrip_1like.
DR Pfam: PF00023; ank; 3.
DR Pfam: PF01412; ArfGAP_1.
DR PRINTS: PR00405; REVINTRACTING.
DR SMART: SM00248; ANK; 1.
DR SMART: SM00105; ArfGAP_1.
DR PROSITE: PS50086; ANK_REPEAT; 1.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE: PS50115; ARFGAP_1.
KW GTPase activation; Repeat; ANK repeat; Zinc-finger; Phosphorylation.
FT DOMAIN 1 124 ARF-GAP.
FT ZN FING 11 34 C4-TYPE.
FT REPEAT 132 161 ANK 1.
FT REPEAT 166 195 ANK 2.
FT REPEAT 199 228 ANK 3.
SQ SEQUENCE 708 AA; 78795 MW; BEA1C5D74182D8CA CRC64;

Query Match
Best Local Similarity 33.3%; Score 5; DB 1; Length 708;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STRIPS 12
DB 535 STRIPS 539

RESULT 150
NTT4_RAT
ID NTT4_RAT STANDARD; PRT; 727 AA.
AC P31662;
DT 01-JUL-1993 (Rel. 26; Created)
DT 01-JUL-1993 (Rel. 26; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Orphan sodium- and chloride-dependent neurotransmitter transporter
DE NTT4.
GN NTT4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=9311444; PubMed=8093354;
RA Liu Q.-R., Mandiyan S., Lopez-Corcuera B., Nelson H., Nelson N.;
RA "A rat brain cDNA encoding the neurotransmitter transporter with an
RA unusual structure."
RT FEBS Lett. 315:114-118(1993).
RL [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94125086; PubMed=8294906;
RA el Mestikawy S., Giros B., Pohl M., Hamon M., Kingsmore S.F.,
RA Seidlin M.F., Caron M.G.;
RT "Characterization of an atypical member of the Na+/Cl(-)-dependent
RT transporter family: chromosomal localization and distribution in
RT GABAergic and glutamatergic neurons in the rat brain."
RL J. Neurochem. 62:445-455(1994).
CC -1- FUNCTION: NOT YET KNOWN, ORPHAN TRANSPORTER.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: FOUND EXCLUSIVELY IN THE CENTRAL NERVOUS
CC SYSTEM AND IS MORE ABUNDANT IN THE CEREBELLUM AND THE CEREBRAL
CC CORTEX.
CC -1- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC FAMILY (SMF).
```


DR InterPro: IPR000008; C2.
 DR InterPro: IPR002149; LRI.
 DR InterPro: IPR003315; RPH3A_effector.
 DR InterPro: IPR001565; Synaptotagmin.
 DR InterPro: IPR000306; Znf_FYVE.
 DR Pfam: PF00168; C2_2.
 DR Pfam: PF02318; RPH3A_effector; 1.
 DR PRINTS: PR00360; C2DOMAIN.
 DR PRINTS: PR00399; SYNAPTOTAGMIN.
 DR SMART: SM00239; C2_2.
 DR PROSITE: PS00499; C2_DOMAIN_1; 2.
 DR PROSITE: PS50004; C2_DOMAIN_2; 2.
 DR PROSITE: PS50178; ZF_FYVE; 1.
 DR Repeat: Synapse; Protein transport; Zinc-finger; 3D-structure.
 KW DOMAIN 1 277 RAB_P25/SMG_P25A_BINDING (GTP GAMMA-S-
 BOUND).
 FT ZN_FING 88 145 FYVE-TYPE.
 FT DOMAIN 280 364 PRO-RICH.
 FT DOMAIN 384 488 C2 DOMAIN 1.
 FT DOMAIN 542 645 C2 DOMAIN 2.
 SQ SEQUENCE 684 AA; 75832 MM; 05838BC3C7A86444 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 684;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESR 13
 |||||
 DB 188 TPESR 192

RESULT 147
 NOPY_YEAST
 ID NOPY_YEAST STANDARD: PRT: 685 AA.

AC P37838; 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Nucleolar protein NOP4 (Nucleolar protein NOP77).
 GN NOP4 OR NOP77 OR YPL043W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RP [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94313993; PubMed=8039505;
 RA Sun C., Woolford J.L., Jr.;
 RT "The yeast NOP4 gene product is an essential nucleolar protein
 RT required for pre-rRNA processing and accumulation of 60S ribosomal
 RT subunits.";
 RL EMO J. 13:3127-3135(1994).
 RN [2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=94313994; PubMed=8039506;
 RA Berges T., Petfalaski E., Tollervey D., Hurt E.C.;
 RT "Synthetic lethality with fibrillarin identifies NOP77p, a nucleolar
 RT protein required for pre-rRNA processing and modification.";
 RL EMO J. 13:3136-3148(1994).
 RN [3]
 RN SEQUENCE FROM N.A.
 RP STRAIN=S288C / AB972;
 RC MEDLINE=97313271; PubMed=9169875;
 RA Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansoorge W.,
 RA Arujo R., Aparicio A., Barrell B.G., Badcock K., Benes V.,
 RA Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,
 RA Chung E., Churcher C.M., Coster F., Davis R.W.,
 RA Dietrich F.S., Delius H., DiPaolo T., Dubois E., Duesterhoeft A.,
 RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,
 RA Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.,
 RA Hunicke-Smith S., Hyman R., Johnston M., Kalman S., Kleene K.,
 RA Komp C., Kurtl O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,
 RA Marathe R., Messenguy F., Mewes H.-W., Mittlepat S., Moestl D.,

RA Mueller-Auer S., Namath A., Nentwich U., Oefner P., Pearson D.,
 RA Petel F.X., Pohl T.M., Purnelle D., Schafer M., Scharfe M.,
 RA Scherens B., Schramm S., Schroeder M., Sdicu A.M., Tettelin H.,
 RA Urestrazaru L.A., Ushinsky S., Vierendeels F., Vissers S., Voss H.,
 RA Walsh S.V., Wambutt R., Wang Y., Wedler E., Wedler H., Winnett E.,
 RA Zhong W.W., Zollner A., Vo D.H., Hant J.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI.";
 RL Nature 387:103-105(1997).
 CC -1 FUNCTION: REQUIRED FOR 60S RIBOSOMAL SUBUNIT SYNTHESIS. PROBABLY
 CC INVOLVED IN THE PROCESSING OF 27S RNA TO PRODUCE MATURE 25S RNA.
 CC -1 SUBUNIT: INTERACTS WITH NOP1.
 CC -1 SUBCELLULAR LOCATION: Nucleolus; nucleolar.
 CC -1 SIMILARITY: CONTAINS 4 RNA RECOGNITION MOTIFS (RRM).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: U02598; AAA20590.1; -;
 DR EMBL: X76245; CA53824.1; -;
 DR EMBL: U44030; AAB68177.1; -;
 DR PIR: S46365; S46365.
 DR PIR: S45178; S45178.
 DR HSSP: P06651; 1HA1.
 DR SGD: S0005964; NOP4.
 DR InterPro: IPR000504; RNA_rec_mot.
 DR Pfam: PF00076; rrm; 4.
 DR SMART: SM00360; RRM; 4.
 DR PROSITE: PS50102; RRM; 4.
 DR PROSITE: PS00030; RRM_RNP_1; 3.
 KW Ribosome biogenesis; Nucleolus; RNA-binding; rRNA processing;
 KW Repeat.
 FT DOMAIN 26 103 RNA-BINDING (RRM) 1.
 FT DOMAIN 147 225 RNA-BINDING (RRM) 2.
 FT DOMAIN 290 383 RNA-BINDING (RRM) 3.
 FT DOMAIN 462 612 RNA-BINDING (RRM) 4.
 FT DOMAIN 241 267 ASP/GLU-RICH (ACIDIC).
 FT VARIANT 308 308 P -> A.
 SQ SEQUENCE 685 AA; 77825 MM; 61261815EA3DE5C CRC64;

Query Match 33.3%; Score 5; DB 1; Length 685;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STRES 12
 |||||
 DB 117 STRES 121

RESULT 148
 HS83_LEIIN
 ID HS83_LEIIN STANDARD: PRT: 701 AA.
 AC Q25293;
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Heat shock protein 83-1 (HSP 83).
 DE HSP83-1.
 OS Leishmania infantum.
 GN Leishmania infantum.
 OC Eukaryota; Eumetazoa; Kinetoplastida; Trypanosomatidae; Leishmanina.
 OX NCBI_TaxID=56711;
 RP [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LEM 75 / Zymodeme 1;
 RX MEDLINE=97126362; PubMed=8971277;
 RA Angel S.O., Requena J.M., Soto M., Criado D., Alonso C.;
 RT "During canine leishmaniasis a protein belonging to the 83-kDa
 RT heat-shock protein family elicits a strong humoral response.";
 RL Acta Trop. 62:45-56(1996).

DR Pfam: PF00515; TPR; 2.
 KW Nuclear protein; Sperm; Coiled coil; Antigen.
 FT DOMAIN 16 57 COILED COIL (POTENTIAL).
 FT DOMAIN 351 386 COILED COIL (POTENTIAL).
 FT DOMAIN 490 558 COILED COIL (POTENTIAL).
 FT DOMAIN 645 661 COILED COIL (POTENTIAL).
 FT DOMAIN 16 26 GUU-RICH (ACIDIC).
 FT DOMAIN 354 404 GUU-RICH (ACIDIC).
 FT DOMAIN 609 615 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 SO SEQUENCE 680 AA; 73524 MW; CC3CA5692CB16F60 CRC64;
 Query Match 33.3%; Score 5; DB 1; Length 680;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 11 ESRFA 15
 DB 660 ESRFA 664
 RESULT 145
 RP3A_MOUSE STANDARD; PRT: 681 AA.
 ID RP3A_MOUSE
 AC P47708;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Rabphilin-3A.
 GN RPH3A.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RA Inaaki N.;
 RL Submitted (NOV-2001) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1-606 FROM N.A.
 RC STRAIN=C57BL/6;
 RX MEDLINE=95122445; PubMed=7822236;
 RA Inaaki N., Mizuta M., Selno S.;
 RT "Cloning of a mouse Rabphilin-3A expressed in hormone-secreting
 cells.";
 RL J. Biochem. 116:239-242(1994).
 CC -1- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED WITH RAS-RELATED
 CC PROTEIN RAB-3A IN SYNAPTIC VESICLE TRAFFIC AND/OR SYNAPTIC
 CC VESICLE FUSION. COULD PLAY A ROLE IN NEUROTRANSMITTER RELEASE BY
 CC REGULATING MEMBRANE FLOW IN THE NERVE TERMINAL.
 CC -1- SUBUNIT: MONOMER.
 CC -1- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN BRAIN.
 CC -1- SIMILARITY: CONTAINS 2 C2 DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 FYVE-TYPE ZINC FINGER.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D29965; BAA06231.2; -
 DR HSSP: P47709; 1ZBD.
 DR MGD: MGI:102786; Rph3a.
 DR InterPro: IPR000008; C2.
 DR InterPro: IPR002149; LRI.
 DR InterPro: IPR00315; RPH3A_effector.
 DR InterPro: IPR001565; Synaptotagmin.
 DR InterPro: IPR000306; Znf_FYVE.
 DR Pfam: PF00168; C2; 2.
 DR Pfam: PF02318; RPH3A_effector; 1.

DR PRINTS; PR00360; C2DOMAIN.
 DR PRINTS; PR00399; SYNAPTOTAGMN.
 DR SMART; SM00239; C2; 1.
 DR PROSITE; PS00499; C2_DOMAIN_1; 2.
 DR PROSITE; PS50004; C2_DOMAIN_2; 2.
 DR PROSITE; PS50178; ZF_FYVE; 1.
 KW Repeat; Synapse; Protein transport; Zinc-finger.
 FT DOMAIN 1 274 RAB P25/SMG P25A BINDING (GTP GAMMA-S-
 FT ZN-FING 88 145 BOUND).
 FT DOMAIN 277 361 FYVE-TYPE.
 FT DOMAIN 381 485 PRO-RICH.
 FT DOMAIN 539 642 C2 DOMAIN 1.
 FT DOMAIN 539 642 C2 DOMAIN 2.
 SO SEQUENCE 681 AA; 75489 MW; D09F8DBD2CB271E CRC64;
 Query Match 33.3%; Score 5; DB 1; Length 681;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 9 TPESR 13
 DB 188 TPESR 192
 RESULT 146
 RP3A_RAT STANDARD; PRT: 684 AA.
 ID RP3A_RAT
 AC P47709;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Rabphilin-3A.
 GN RPH3A.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95033210; PubMed=7946335;
 RA Li C., Takei K., Geppert M., Daniell L., Stanulus K., Chapman E.R.,
 RA Jahn R., de Camilli P., Südhof T.C.;
 RT "Synaptic targeting of rabphilin-3A, a synaptic vesicle
 RT Ca2+/phospholipid-binding protein, depends on rab3A/3C.";
 RL Neuron 13:885-898(1994).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 44-167 IN COMPLEX WITH RAB3A.
 RC TISSUE=Brain;
 RX MEDLINE=99148269; PubMed=10025402;
 RA Ostermeier C., Brunger A.T.;
 RT "Structural basis of Rab effector specificity: crystal structure of
 RT the small G protein Rab3A complexed with the effector domain of
 RT rabphilin-3A.";
 RL Cell 96:363-374(1999).
 CC -1- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED WITH RAS-RELATED
 CC PROTEIN RAB-3A IN SYNAPTIC VESICLE TRAFFIC AND/OR SYNAPTIC
 CC VESICLE FUSION. COULD PLAY A ROLE IN NEUROTRANSMITTER RELEASE BY
 CC REGULATING MEMBRANE FLOW IN THE NERVE TERMINAL.
 CC -1- SUBUNIT: MONOMER.
 CC -1- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN BRAIN.
 CC -1- SIMILARITY: CONTAINS 2 C2 DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 FYVE-TYPE ZINC FINGER.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U12571; AAA62662.1; -
 DR PDB: 1ZBD; 12-APR-99.

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X13460; CAA31808.1; -
DR PIR; S01786; S01786.
DR HSSP; P79134; IAVC.
DR MGI; MGI:88255; Annex6.
DR InterPro; IPR001464; Annexin.
DR Pfam; PF00191; annexin; 8.
DR PRINTS; PR00196; ANNEXIN.
DR ProDom; PD000143; Annexin; 8.
DR SMART; SM00335; ANX; 8.
DR PROSITE; PS00223; ANNEXIN; 8.
KW Annexin; Calcium/phospholipid-binding; Repeat.
FT INIT_MET 0
FT REPEAT 28 88 ANNEXIN 1.
FT REPEAT 100 160 ANNEXIN 2.
FT REPEAT 184 244 ANNEXIN 3.
FT REPEAT 259 319 ANNEXIN 4.
FT REPEAT 371 431 ANNEXIN 5.
FT REPEAT 443 503 ANNEXIN 6.
FT REPEAT 532 592 ANNEXIN 7.
FT REPEAT 607 667 ANNEXIN 8.
SQ SEQUENCE 672 AA; 75755 MW; 2D85C11DD235FC76 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 672;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
DB 249 RSTPE 253

RESULT 143
ANX6_RAT STANDARD; PRT; 672 AA.
AC P48037;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Annexin VI (lipocortin VI) (P68) (P70) (Protein III) (Chromobindin 20)
DE (67 kDa callectrin) (calphobindin-II) (CPB-II) (Calcium-binding
DE protein CTR 65/67).
GN ANX6 OR ANX6.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=95331313; PubMed=7607247;
RA Fan H., Josic D., Lim Y.P., Reutter W.;
RT "cDNA cloning and tissue-specific regulation of expression of rat
RT calcium-binding protein 65/67. Identification as a homologue of
RT annexin VI."
RL Eur. J. Biochem. 230:741-751(1995).
CC -1- FUNCTION: MAY ASSOCIATE WITH CD21. MAY REGULATE THE RELEASE OF
CC CA(2+) FROM INTRACELLULAR STORES.
CC -1- DOMAIN: A pair of annexin repeats may form one binding site for
CC calcium and phospholipid.
CC -1- MISCELLANEOUS: SEEMS TO BIND ONE CALCIUM ION WITH HIGH AFFINITY.
CC -1- SIMILARITY: BELONGS TO THE ANNEXIN FAMILY.
CC -1- SIMILARITY: CONTAINS 8 ANNEXIN REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X86086; CAA60040.1; -
DR HSSP; P79134; IAVC.
DR InterPro; IPR001464; Annexin.
DR Pfam; PF00191; annexin; 8.
DR PRINTS; PR00196; ANNEXIN.
DR ProDom; PD000143; Annexin; 8.
DR SMART; SM00335; ANX; 8.
DR PROSITE; PS00223; ANNEXIN; 7.
KW Annexin; Calcium/phospholipid-binding; Repeat.
FT INIT_MET 0
FT REPEAT 28 88 ANNEXIN 1.
FT REPEAT 100 160 ANNEXIN 2.
FT REPEAT 184 244 ANNEXIN 3.
FT REPEAT 259 319 ANNEXIN 4.
FT REPEAT 371 431 ANNEXIN 5.
FT REPEAT 443 503 ANNEXIN 6.
FT REPEAT 532 592 ANNEXIN 7.
FT REPEAT 607 667 ANNEXIN 8.
SQ SEQUENCE 672 AA; 75622 MW; BB8E798A5CAB1511 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 672;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
DB 249 RSTPE 253

RESULT 144
NASP_RABIT STANDARD; PRT; 680 AA.
AC P27123;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Nuclear autoantigenic sperm protein (NASP).
GN NASP.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
CX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand white; TISSUE=Testis;
RX MEDLINE=91145522; PubMed=2289010;
RA Welch J.E., Zimmerman L.J., Joseph D.R., O'Rand M.G.;
RT "Characterization of a sperm-specific nuclear autoantigenic protein.
RT I. Complete sequence and homology with the Xenopus protein, NI/N2."
RL Biol. Reprod. 43:559-568(1990).
CC -1- FUNCTION: MAY PLAY A ROLE IN REGULATING THE EARLY EVENTS OF
CC SPERMATOGENESIS BY BINDING NEWLY SYNTHESIZED HISTONES VARIANTS AND
CC TRANSPORTING THEM TO THE NUCLEUS.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: TESTIS- AND SPERM-SPECIFIC PROTEIN.
CC -1- SIMILARITY: WITH XENOPUS HISTONE-BINDING PROTEIN NI/N2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M37893; AAA31423.1; -
DR PIR; A43800; A43800.
DR InterPro; IPR001440; TPR.

ANX6_HUMAN STANDARD; PRT; 672 AA.
 ID ANX6_HUMAN
 AC P08133;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Annexin VI (Lipocortin VI) (P68) (Protein III) (Chromobindin 20)
 DE (67 kDa calelectrin) (Calphobindin-II) (CPB-II).
 GN ANX6 OR ANX6.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RX MEDLINE=88196081; PubMed=3258820;
 RA Crompton M.R., Owens R.J., Totly N.F., Moss S.E., Waterfield M.D.,
 RA Crumpton M.J.;
 RT "Primary structure of the human, membrane-associated Ca²⁺-binding
 RL protein p68 a novel member of a protein family.";
 RN EMBO J. 7:21-27(1988).
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=88124902; PubMed=2963335;
 RA Suedhof T.C., Slaughter C.A., Leznicki I., Barjon P., Reynolds G.A.;
 RT "Human 67-kDa calelectrin contains a duplication of four repeats
 RT found in 35-kDa lipocortins.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:664-668(1988).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89380132; PubMed=2528541;
 RA Iwasaki A., Suda M., Watanabe M., Nakao H., Hattori Y., Nagoya T.,
 RA Saito Y., Shidara Y., Maki M.;
 RT "Structure and expression of cDNA for calphobindin II, a human
 RT placental coagulation inhibitor.";
 RL J. Biochem. 106:43-49(1989).
 RN [4]
 RP SEQUENCE.
 RX MEDLINE=90236978; PubMed=2139657;
 RA Yoshizaki H., Mizoguchi T., Arai K., Shiratsuchi M., Shidara Y.,
 RA Maki M.;
 RT "Structure and properties of calphobindin II, an anticoagulant
 RT protein from human placenta.";
 RL J. Biochem. 107:43-50(1990).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE=uterus;
 RA Strausberg R.;
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RX X-RAY CRYSTALLOGRAPHY (3.2 ANGSTROMS).
 RX MEDLINE=96326697; PubMed=8709144;
 RA Benz J., Bergner A., Hofmann A., Demange P., Goettig P., Liemann S.,
 RA Huber R., Voges D.;
 RT "The structure of recombinant human annexin VI in crystals and
 RT membrane-bound.";
 RL J. Mol. Biol. 260:638-643(1996).
 CC -1- FUNCTION: MAY ASSOCIATE WITH CD21. MAY REGULATE THE RELEASE OF
 CC CA(2+) FROM INTRACELLULAR STORES.
 CC -1- INDUCTION: BY EBV.
 CC -1- DOMAIN: A pair of annexin repeats may form one binding site for
 CC calcium and phospholipid.
 CC -1- PTM: PHOSPHORYLATED IN RESPONSE TO GROWTH FACTOR STIMULATION.
 CC -1- MISCELLANEOUS: SEEMS TO BIND ONE CALCIUM ION WITH HIGH AFFINITY.
 CC -1- SIMILARITY: BELONGS TO THE ANNEXIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 8 ANNEXIN REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

or send an email to license@isb-sib.ch.
 CC -----
 CC EMBL: D00510; BAA00400.1; -;
 CC DR EMBL: Y00097; CAA68286.1; -;
 CC DR EMBL: J03578; AAB56556.1; -;
 CC DR EMBL: BC017046; AAH17046.1; -;
 CC DR PIR: J00032; AQH068.
 CC DR HSSP: P79134; IAVC.
 CC DR Genew: HGNC:544; ANX6.
 CC MIM: 114070; -;
 CC DR InterPro: IPR001464; Annexin.
 CC DR Pfam: PF00191; annexin; 8.
 CC DR PRINTS: PR00196; ANNEXIN.
 CC DR PRODOM: PD000143; Annexin; 8.
 CC DR SMART: SM00335; ANX; 8.
 CC DR PROSITE: PS00223; ANNEXIN; 8.
 CC KW Annexin; Calcium/phospholipid-binding; Repeat; Acetylation;
 CC Phosphorylation.
 CC MW
 CC FT INIT_MET 0 0
 CC FT REPEAT 28 88 ANNEXIN 1.
 CC FT REPEAT 100 160 ANNEXIN 2.
 CC FT REPEAT 184 244 ANNEXIN 3.
 CC FT REPEAT 259 319 ANNEXIN 4.
 CC FT REPEAT 371 431 ANNEXIN 5.
 CC FT REPEAT 443 503 ANNEXIN 6.
 CC FT REPEAT 532 592 ANNEXIN 7.
 CC FT REPEAT 607 667 ANNEXIN 8.
 CC FT MOD_RES 1 1 ACETYLATION.
 CC FT CONFLICT 225 226 IE -> MK (IN REF. 2).
 CC FT CONFLICT 554 554 S -> T (IN REF. 2).
 CC FT CONFLICT 618 618 E -> D (IN REF. 1).
 CC SQ SEQUENCE 672 AA; 75742 MW; 2829237029BDIDCB CRC64;
 CC -----
 CC Query Match 33.3%; Score 5; DB 1; Length 672;
 CC Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC -----
 CC QY 7 RSTPE 11
 CC DB 249 RSTPE 253
 CC -----
 CC RESULT 142
 CC ANX6_MOUSE
 CC ID ANX6_MOUSE STANDARD; PRT; 672 AA.
 AC P14824;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Annexin VI (Lipocortin VI) (P68) (Protein III) (Chromobindin 20)
 DE (67 kDa calelectrin) (Calphobindin-II) (CPB-II).
 GN ANX6 OR ANX6.
 OS Mus musculus (mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BAIR/c;
 RX MEDLINE=89030687; PubMed=2972541;
 RA Moss S.E., Crompton M.R., Crumpton M.J.;
 RT "Molecular cloning of murine p68, a Ca²⁺-binding protein of the
 RT lipocortin family.";
 RL Eur. J. Biochem. 177:21-27(1988).
 CC -1- FUNCTION: MAY ASSOCIATE WITH CD21. MAY REGULATE THE RELEASE OF
 CC CA(2+) FROM INTRACELLULAR STORES.
 CC -1- DOMAIN: A pair of annexin repeats may form one binding site for
 CC calcium and phospholipid.
 CC -1- MISCELLANEOUS: SEEMS TO BIND ONE CALCIUM ION WITH HIGH AFFINITY.
 CC -1- SIMILARITY: BELONGS TO THE ANNEXIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 8 ANNEXIN REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```
CC -----
DR EMBL: AL583918; CAC30056.1; -.
DR Leptoma; ML0548; -.
DR InterPro: IPR005259; P1A.
DR TIGRfams; TIGR00595; P1A; 1.
KW DNA replication; DNA-binding; Helicase; Primosome; Zinc-finger;
FT ZN-FING 371 383 C4-TYPE (POTENTIAL).
FT ZN-FING 399 414 C4-TYPE (POTENTIAL).
SQ SEQUENCE 651 AA; 69923 MW; 06CFE91D208C7043 CRC64;

Query Match
Best Local Similarity 100.0%; Score 5; DB 1; Length 651;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 11 ESRAA 15
Db 141 ESRAA 145

RESULT 139
MOTL_NEIMA STANDARD; PRT; 658 AA.
AC 09JTS2;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE DNA mismatch repair protein mutL.
GN MUTL OR NMA1655
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=22491 / Serogroup A / Serotype 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagsels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrell B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
meningitidis 22491."
RL Nature 404:502-506(2000).
CC -1- FUNCTION: This protein is involved in the repair of mismatches in
CC DNA. It is required for dam-dependent methyl-directed DNA mismatch
CC repair. May act as a "molecular matchmaker", a protein that
CC promotes the formation of a stable complex between two or more
CC DNA-binding proteins in an ATP-dependent manner without itself
CC being part of a final effector complex (By similarity).
CC -1- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MOTL/HEXB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL162756; CAB84883.1; -.
DR HSSP: P23367; IBKN.
DR InterPro: IPR003594; ATPbind_ATPase.
DR InterPro: IPR002099; DNA_mis_repair.
DR Pfam: PF01119; DNA_mis_repair; 1.
DR Pfam: PF02518; HATPase_C; 1.
DR TIGRfams; TIGR00585; mutL; 1.
DR PROSITE; PS00058; DNA_MISMATCH_REPAIR_1; FALSE_NEG.
KW DNA repair; Complete proteome.
SQ SEQUENCE 658 AA; 71841 MW; 76B33C9628F63506 CRC64;

Query Match
33.3%; Score 5; DB 1; Length 658;
```

```
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 11 ESRAA 15
Db 408 ESRAA 412

RESULT 140
MOTL_NEIMA STANDARD; PRT; 658 AA.
AC 09JTS2;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE DNA mismatch repair protein mutL.
GN MUTL OR NMA1442.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / Serogroup B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson R.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.R.,
RA Hart D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Clecko A., Parksey D.S., Blair E., Citterone H., Clark E.B.,
RA Cotton M.D., Utterback T.R., Khouli H., Qin H., Yamathevan J.,
RA Gill J., Scarlato V., Maignani V., Pizzo M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
MC58."
RL Science 287:1809-1815(2000).
CC -1- FUNCTION: This protein is involved in the repair of mismatches in
CC DNA. It is required for dam-dependent methyl-directed DNA mismatch
CC repair. May act as a "molecular matchmaker", a protein that
CC promotes the formation of a stable complex between two or more
CC DNA-binding proteins in an ATP-dependent manner without itself
CC being part of a final effector complex (By similarity).
CC -1- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MOTL/HEXB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE002493; AAF41803.1; -.
DR HSSP: P23367; IBKN.
DR TIGR: NMB1442; -.
DR InterPro: IPR003594; ATPbind_ATPase.
DR InterPro: IPR002099; DNA_mis_repair.
DR Pfam: PF01119; DNA_mis_repair; 1.
DR Pfam: PF02518; HATPase_C; 1.
DR TIGRfams; TIGR00585; mutL; 1.
DR PROSITE; PS00058; DNA_MISMATCH_REPAIR_1; FALSE_NEG.
KW DNA repair; Complete proteome.
SQ SEQUENCE 658 AA; 71646 MW; 40BEDC060396DD2F CRC64;

Query Match
33.3%; Score 5; DB 1; Length 658;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 11 ESRAA 15
Db 408 ESRAA 412

RESULT 141
```

```

RA  Ahrens C.H., Russell R.R., Funk C.J., Evans J., Harwood S.,
RA  Rohmann G.F.,"
RT  "The sequence of the Orygia pseudotsugata multinucleocapsid nuclear
RL  polyhedrosis virus genome."
RL  Virology 229:381-399(1997).
CC  -1- FUNCTION: ESSENTIAL FOR VIRULENCE OF BACULOVIRUS OCCUSION
CC  BODIES FOR INSECT LARVAE (BY SIMILARITY).
CC  -----
CC  CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; U75930; AAC59133.1;
DR  Virulence; Viral occlusion body.
SO  SEQUENCE 644 AA; 72653 MW; AEBF901D0FEC9810 CRC64;

QY  11 ESRAA 15
QY  |||||
Db  327 ESRAA 331

RESULT 138
PR1A_MYCLE STANDARD; PRT; 651 AA.
AC  09CCQ3;
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  Putative primosomal protein N' (Replication factor Y).
GN  P1A OR ML0548.
OS  Mycobacterium leprae.
OC  Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC  Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX  NCBI_TaxID=1769;
[1]
RN  SEQUENCE FROM N.A.
RP  STRAIN=TN.
RX  MEDLINE=21128732; PubMed=11234002;
RA  Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA  Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA  Davies R.M., Devlin K., Dutfoy S., Fellwell T., Fraser A., Hamlin N.,
RA  Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA  Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA  Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA  Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA  Barrett B.G.;
RT  "Massive gene decay in the Leprosy bacillus."
RL  Nature 409:1007-1011(2001).
CC  -1- FUNCTION: RECOGNIZES A SPECIFIC HAIRPIN SEQUENCE ON PHX SSDNA;
CC  THIS STRUCTURE IS THEN RECOGNIZED AND BOUND BY PROMPTING PR1A AND
CC  PRIC. FORMATION OF THE PRIMOSOME PROCEEDS WITH THE SUBSEQUENT
CC  ACTIONS OF DNAB, DNAC, DNAT AND PRIMASE. PR1A THEN FUNCTIONS AS A
CC  HELICASE WITHIN THE PRIMOSOME (BY SIMILARITY).
CC  -1- SIMILARITY: BELONGS TO THE HELICASE FAMILY. PR1A SUBFAMILY.
CC  -1- CAUTION: COMPARED TO OTHER BACTERIAL PR1A, IT HAS A VERY DIVERGENT
CC  HELICASE DOMAIN.
CC  -----
CC  CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).

```

CC STRAIN-Sprague-Dawley; TISSUE=Osteosarcoma;
 RX MEDLINE=93094228; PubMed=1334084;
 RA Chan S.D.H., Karpf D.B., Fowlkes M.E., Hooks M., Bradley M.S.,
 RA Vuong V., Bambino T., Liu M.Y.C., Arnaud C.D., Sreewler G.O.,
 RA Nissenson R.A.;
 RT "Two homologs of the Drosophila polarity gene frizzled (fz) are widely
 RT expressed in mammalian tissues.";
 RL J. Biol. Chem. 267:25202-25207(1992).
 RN [2]
 RP COUPLING TO BETA-CATENIN PATHWAY.
 RX MEDLINE=99324245; PubMed=10395342;
 RA Sheldahl L.C., Park M., Malbon C.C., Moon R.T.;
 RT "Protein kinase C is differentially stimulated by Wnt and Frizzled
 RT homologs in a G-protein-dependent manner.";
 RL Curr. Biol. 9:695-698(1999).
 CC -1- FUNCTION: Receptor for Wnt proteins. Most of frizzled receptors
 CC are coupled to the beta-catenin canonical signaling pathway, which
 CC leads to the activation of dishevelled proteins, inhibition of
 CC GSK-3 kinase, nuclear accumulation of beta-catenin and activation
 CC of Wnt target genes. A second signaling pathway involving PKC and
 CC calcium fluxes has been seen for some family members, but it is
 CC not yet clear if it represents a distinct pathway or if it can be
 CC integrated in the canonical pathway, as PKC seems to be required
 CC for Wnt-mediated inactivation of GSK-3 kinase. Both pathways seem
 CC to involve interactions with G-proteins. May be involved in
 CC transduction and intercellular transmission of polarity
 CC information during tissue morphogenesis and/or in differentiated
 CC tissues. Activation by Wnt8 induces expression of beta-catenin
 CC target genes.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Widely expressed. Most abundant in kidney,
 CC liver, uterus, ovary and heart. Lower levels seen in brain and
 CC intestine. Extremely low in calvaria, mammary glands and testis.
 CC -1- DEVELOPMENTAL STAGE: Expressed predominantly in neonatal tissues,
 CC at lower levels in adult.
 CC -1- DOMAIN: Lys-Thr-X-X-Tip motif is involved in the activation of
 CC the Wnt/beta-catenin signaling pathway (By similarity).
 CC -1- DOMAIN: The fz domain is involved in binding with Wnt ligands (By
 CC similarity).
 CC -1- SIMILARITY: BELONGS TO FAMILY FZ/SMO OF G-PROTEIN COUPLED
 CC RECEPTORS.
 CC -1- SIMILARITY: CONTAINS 1 FRIZZLED (FZ) DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: L02529; AAA1173.1; -
 DR InterPro: IPR000539; Frizzled.
 DR InterPro: IPR000024; Fz_domain.
 DR InterPro: IPR000832; GPCR_secretin.
 DR Pfam: PF01392; Fz; 1.
 DR Pfam: PF01534; Frizzled; 1.
 DR PRINTS: PR00489; FRIZZLED.
 DR SMART: SM00063; FRI; 1.
 DR PROSITE: PSS0038; FZ; 1.
 DR PROSITE: PSS0261; G-PROTEIN_RECPT_F2_4; 1.
 KW Multigene family; G-protein coupled receptor; Transmembrane;
 KW developmental protein; Glycoprotein; Signal.
 FT CHAIN 1 68 POTENTIAL.
 FT SIGNAL 1 68 FRIZZLED 1.
 FT DOMAIN 69 641 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 317 337 1 (POTENTIAL).
 FT DOMAIN 338 348 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 349 369 2 (POTENTIAL).
 FT DOMAIN 370 396 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 397 417 3 (POTENTIAL).
 FT DOMAIN 418 439 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 440 460 4 (POTENTIAL).

FT DOMAIN 461 483 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 484 504 5 (POTENTIAL).
 FT DOMAIN 505 530 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 531 551 6 (POTENTIAL).
 FT DOMAIN 552 595 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 596 616 7 (POTENTIAL).
 FT DOMAIN 617 641 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 106 224 FZ.
 FT DOMAIN 85 90 POLY-PRO.
 FT SITE 619 624 LYS-THR-X-X-TRP MOTIF.
 FT SITE 639 641 PDZ-BINDING.
 FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 225 225 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 641 AA; 71027 MM; D82EC113E81B886 CRC64;
 Query Match 33.3%; Score 5; DB 1; Length 641;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 11 ESRAA 15
 Db 9 ESRAA 13
 ID FZD1_MOUSE STANDARD; PRT; 642 AA.
 AC 070421; 008974;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Frizzled 1 precursor (Frizzled-1) (Fz-1) (Mfz1).
 GN FZD1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Embryo;
 RX MEDLINE=21233588; PubMed=11334716;
 RA Xu L., Tan L., Goldring M.B., Olsen B.R., Li Y.;
 RT "Expression of frizzled genes in mouse costochondral chondrocytes.";
 RL Matrix Biol. 20:147-151(2001).
 RN [2]
 RP SEQUENCE OF 302-376 FROM N.A.
 RC TISSUE=Prostate;
 RA Johnson M.A., Greenberg N.M.;
 RT Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Receptor for Wnt proteins. Most of frizzled receptors
 CC are coupled to the beta-catenin canonical signaling pathway, which
 CC leads to the activation of dishevelled proteins, inhibition of
 CC GSK-3 kinase, nuclear accumulation of beta-catenin and activation
 CC of Wnt target genes. A second signaling pathway involving PKC and
 CC calcium fluxes has been seen for some family members, but it is
 CC not yet clear if it represents a distinct pathway or if it can be
 CC integrated in the canonical pathway, as PKC seems to be required
 CC for Wnt-mediated inactivation of GSK-3 kinase. Both pathways seem
 CC to involve interactions with G-proteins. May be involved in
 CC transduction and intercellular transmission of polarity
 CC information during tissue morphogenesis and/or in differentiated
 CC tissues.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Expressed in chondrocytes.
 CC -1- DOMAIN: Lys-Thr-X-X-Tip motif is involved in the activation of
 CC the Wnt/beta-catenin signaling pathway (By similarity).
 CC -1- DOMAIN: The fz domain is involved in binding with Wnt ligands (By
 CC similarity).
 CC -1- SIMILARITY: BELONGS TO FAMILY FZ/SMO OF G-PROTEIN COUPLED
 CC RECEPTORS.
 CC -1- SIMILARITY: CONTAINS 1 FRIZZLED (FZ) DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003598; Ig_C2.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00047; Ig_1.
DR SMART: SM00181; EGF_1.
DR SMART: SM00408; ICG2_1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; FALSE_NEG.
KW EGF-like domain; Glycoprotein; Immunoglobulin domain; Growth factor;
KW Developmental protein; Alternative splicing; Signal.
FT SIGNAL 1 42
FT CHAIN 1 622
FT DOMAIN 470 537 VEIN PROTEIN.
FT DOMAIN 561 599 IG-LIKE C2-TYPE DOMAIN.
FT DOMAIN 43 53 EGF-LIKE.
FT DOMAIN 87 96 POLY-GLN.
FT DOMAIN 149 157 POLY-SER.
FT DOMAIN 283 291 POLY-GLN.
FT DOMAIN 299 304 POLY-GLN.
FT DISULFID 477 530 BY SIMILARITY.
FT CARBOHYD 76 76 N-LINKED (GLCNAC . . .) (POTENTIAL).
FT CARBOHYD 210 210 N-LINKED (GLCNAC . . .) (POTENTIAL).
FT CARBOHYD 231 231 N-LINKED (GLCNAC . . .) (POTENTIAL).
FT CARBOHYD 349 349 N-LINKED (GLCNAC . . .) (POTENTIAL).
FT CARBOHYD 380 380 N-LINKED (GLCNAC . . .) (POTENTIAL).
FT CARBOHYD 423 423 N-LINKED (GLCNAC . . .) (POTENTIAL).
FT CARBOHYD 448 448 N-LINKED (GLCNAC . . .) (POTENTIAL).
FT CARBOHYD 520 520 N-LINKED (GLCNAC . . .) (POTENTIAL).
FT CARBOHYD 573 573 N-LINKED (GLCNAC . . .) (POTENTIAL).
FT VARSPLIC 608 622 FVAIIGQIHITLNDY -> SSPECKNTYGGY (IN ISOFORM 2).
SQ SEQUENCE 622 AA; 71569 MW; 910C67F0A1E76D45 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 622;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRA 14
Db 201 PESRA 205

RESULT 133
V70K_TYMW STANDARD; PRT; 628 AA.
AC P10357;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE 69 kDa protein.
OS Turnip yellow mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tymovirus.
OX NCBI_TaxID=12154;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88289359; PubMed=3399388;
RA Morch M.D.; Boyer J.C.; Haenni A.L.;
RT "Overlapping open reading frames revealed by complete nucleotide
sequencing of turnip yellow mosaic virus genomic RNA";
RL Nucleic Acids Res. 16:6157-6173(1988).
CC -1- FUNCTION: NOT KNOWN.
CC -1- SIMILARITY: TO 65 TO 70 KDa PROTEIN FROM OTHER TYMOVIRUSES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC EMBL; X07441; CAA30321.1; ALT_SEQ.

DR PIR: S01955; S01955.
DR InterPro: IPR004935; TyMo_45_70kDa.
DR Pfam: PF03251; TyMo_45kd_70kd; 1.
SQ SEQUENCE 628 AA; 69195 MW; 9B01CE5ADCEAC77 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 628;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESR 13
Db 472 TPESR 476

RESULT 134
V70K_TYMW STANDARD; PRT; 628 AA.
AC P28478;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE 69 kDa protein.
OS Turnip yellow mosaic virus (isolate TYMC).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tymovirus.
OX NCBI_TaxID=31751;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92119261; PubMed=1731998;
RA Dreher T.W.; Branson K.L.;
RT "Genomic RNA sequence of turnip yellow mosaic virus isolate TYMC, a
CDNA-based clone with verified infectivity";
RL Plant Mol. Biol. 18:403-406(1992).
CC -1- FUNCTION: NOT KNOWN.
CC -1- SIMILARITY: TO 65 TO 70 KDa PROTEIN FROM OTHER TYMOVIRUSES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC EMBL; X16378; CAA34414.1; -.
DR PIR: S19150; S19150.
DR InterPro: IPR004935; TyMo_45_70kDa.
DR Pfam: PF03251; TyMo_45kd_70kd; 1.
SQ SEQUENCE 628 AA; 69057 MW; 0DF0C64E8BCBC6DC CRC64;

Query Match 33.3%; Score 5; DB 1; Length 628;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESR 13
Db 472 TPESR 476

RESULT 135
FZD1_RAT STANDARD; PRT; 641 AA.
AC Q08463;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Frlzzled 1 precursor (Frlzzled-1) (Fz-1) (Fz1).
GN FZD1.
OS Rattus norvegicus (Rat).
CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.

DR Prodom: PD000006; ABC_transporter: 2.
 DR PROSITE; PS00211: ABC_TRANSPORTER: 1.
 KW Hypothetical protein: ATP-binding; Transport; Repeat.
 FT NP_BIND 114 121 ATP (POTENTIAL).
 FT NP_BIND 428 435 ATP (POTENTIAL).
 SQ SEQUENCE 610 AA; 68377 MW; 0E5DB4A3305B4F0 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 610;
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRAA 15
 |||||
 Db 203 ESRAA 207

RESULT 131
 ID IF4B_HUMAN STANDARD; PRT; 611 AA.
 AC P23588;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Eukaryotic translation initiation factor 4B (eIF-4B).
 GN EIF4B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 RX MEDLINE=90360989; PubMed=2390971;
 RA Milburn S.C., Hershey J.W.B., Davies M.V., Kelleher K., Kaufman R.J.;
 RT "Cloning and expression of eukaryotic initiation factor 4B cDNA;
 RT sequence determination identifies a common RNA recognition motif.";
 RL EMBO J. 9:2783-2790(1990).
 RN [2]
 RP CHARACTERIZATION.
 RX MEDLINE=94187701; PubMed=8139536;
 RA Mehot N., Pause A., Hershey J.W., Sonenberg N.;
 RT "The translation initiation factor eIF-4B contains an RNA-binding
 RT region that is distinct and independent from its ribonucleoprotein
 RT consensus sequence.";
 RL Mol. Cell. Biol. 14:2307-2316(1994).
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE=96413282; PubMed=8616444;
 RA Mehot N., Song M.S., Sonenberg N.;
 RT "A region rich in aspartic acid, arginine, tyrosine, and glycine
 RT (DRYG) mediates eukaryotic initiation factor 4B (eIF4B) self-
 RT association and interaction with eIF3.";
 RL Mol. Cell. Biol. 16:5328-5334(1996).
 CC -1- FUNCTION: REQUIRED FOR THE BINDING OF MRNA TO RIBOSOMES. FUNCTIONS
 CC IN CLOSE ASSOCIATION WITH EIF4-F AND EIF4-A. BINDS NEAR THE 5'-
 CC TERMINAL CAP OF MRNA IN PRESENCE OF EIF-4F AND ATP. PROMOTES THE
 CC APPASE ACTIVITY AND THE ATP-DEPENDENT RNA UNWINDING ACTIVITY OF
 CC BOTH EIF4-A AND EIF4-F.
 CC -1- SUBUNIT: SELF-ASSOCIATES AND INTERACTS WITH EIF3 P170 SUBUNIT.
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X55733; CA39265.1; -
 DR PIR; S12566; S12566.
 DR Genew; HGNC:3285; EIF4B.
 DR MIM; 603928; -
 DR InterPro; IPR000504; RNA_rec_mot.

DR Pfam; PF00076; rrm; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS50102; RRM; 1.
 DR PROSITE; PS00030; RRM_NRP_1; 1.
 KW Protein biosynthesis; Initiation factor; RNA-binding.
 FT DOMAIN 96 173 RNA-BINDING (RRM).
 FT DOMAIN 164 356 DRIG.
 SQ SEQUENCE 611 AA; 69224 MW; 5EAD0891694D00D9 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 611;
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPE 11
 |||||
 Db 339 RSTPE 343

RESULT 132
 ID VEIN_DROME STANDARD; PRT; 622 AA.
 AC 094918;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Vein protein precursor (Epidermal growth factor-like protein)
 DE (defective dorsal discs protein).
 GN VN OR DDP.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Imaginal disks, and Embryo;
 RX MEDLINE=96421972; PubMed=8824589;
 RA Schnepp B.C., Grumbly G.B., Donaldson T.D., Simcox A.A.;
 RT "Vein is a novel component in the Drosophila epidermal growth factor
 RT receptor pathway with similarity to the neuregulins.";
 RL Genes Dev. 10:2302-2313(1996).
 CC -1- FUNCTION: LIGAND FOR THE EGFR RECEPTOR. SEEMS TO PLAY A ROLE IN
 CC THE GLOBAL PROLIFERATION OF WING DISC CELLS AND THE LARVAL
 CC PATTERNING. SHOWS A STRONG SYNERGISTIC GENETIC INTERACTION WITH
 CC SPI, SUGGESTING A MOLECULAR INTERDEPENDENCE. REQUIRED FOR THE
 CC DEVELOPMENT OF INTERVEIN CELLS.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN BLASTODERM EMBRYOS IN TWO
 CC VENTROLATERAL STRIPES THAT ARE BROUGHT TO THE MIDLINE AS
 CC GASTRULATION PROCEEDS. IN THE GERM-BAND RETRACTION STAGE,
 CC EXPRESSION IS SEEN IN THE CNS AND EPIDERMIS. AT LATE BLASTODERM,
 CC EXPRESSION IS LOCALIZED IN THE ANLAGEN OF THE AMNIOTEROSA.
 CC EXPRESSION IN THE HEAD, CYCLOLABRUM, MAXILLARY AND LABIAL LOBS,
 CC AND AROUND THE STOMODEUM THROUGHOUT EMBRYO DEVELOPMENT. IN LATE
 CC EMBRYOS, EXPRESSION DECAYS IN ALL ECTODERMAL CELLS AND APPEARS IN
 CC THE SEGMENTAL MUSCLES AND THE GUT WALL. IN THE LARVA, EXPRESSION
 CC OCCURS IN THE DORSAL METATHORACIC DISC, THE EYE-ANTENNAL DISC AND
 CC THE VENTRAL THORACIC DISC. FOUND IN THE INTERVEIN IN THE PUPA.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U67935; AAC47293.1; -
 DR FLYBase; FBgn0003984; vn.

```

OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=391;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96219094; PubMed=8829547;
RA Maruta K., Hattori K., Nakada T., Kubota M., Sugimoto T., Kurimoto M.;
RT "Cloning and sequencing of trehalose biosynthesis genes from Rhizobium
  sp. M-11."
RL Biosci. Biotechnol. Blochem. 60:717-720(1996).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1->4)-D-glucosidic
  linkage in 4-alpha-D-(1->4)-alpha-D-glucanoseyl(n) trehalose to
  yield trehalose and alpha-(1->4)-D-glucan.
CC -1- PATHWAY: Trehalose biosynthesis.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
  KNOWN AS THE ALPHA-AMYLASE FAMILY
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: D78001; BAA1187.1; -.
CC InterPro: IPR000461; Alpha_amylase.
CC InterPro: IPR000493; Isoamylase_N.
CC Pfam: PF00128; alpha-amylase; 1.
CC DR Pfam: PF02922; Isoamylase_N; 1.
CC KW Hydrolyase; Glycosidase.
CC FT ACT_SITE 265 265
CC SQ SEQUENCE 596 AA; 65262 MW; 73EA80AE0534DCCD CRC64;
-----
BY SIMILARITY.
Query Match 33.3%; Score 5; DB 1; Length 596;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 RSTPE 11
ID 11111
DB 518 RSTPE 522
-----
RESULT 129
SYD_SYNX3 STANDARD; PRT; 599 AA.
AC P73851;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--tRNA ligase)
DE (AsPRS).
GN ASPs OR SLR1720.
OS Synchocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synchocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
  Miyajima N., Hikosawa M., Sugiyura A., Sasamoto S., Kimura T.,
  Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Nairo K.,
  Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,
  Yamada M., Yasuda M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
  Synchocystis sp. strain PCC6803. II. Sequence determination of the
  entire genome and assignment of potential protein-coding regions."
RT DNA Res. 3:109-136(1996).
CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +
  dihydrophosphate + L-aspartyl-tRNA(Asp).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.
-----

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: D90910; BAA17910.1; -.
CC DR HSP; P36419; 1EFM.
CC DR InterPro: IPR002106; ATRNA_ligaseII.
CC DR InterPro: IPR004524; Asps_dact.
CC DR InterPro: IPR004115; GAD_dom.
CC DR InterPro: IPR004364; tRNA-synt_2.
CC DR InterPro: IPR002312; tRNA-synt_2.
CC DR InterPro: IPR004365; tRNA-antl.
CC DR Pfam: PF00152; tRNA-synt_2; 2.
CC DR Pfam: PF01336; tRNA-antl; 1.
CC DR Pfam: PF02938; GAD; 1.
CC DR PRINTS: PR01042; TRNASYNTASP.
CC DR TIGRfam: TIGR00459; asps_dact; 1.
CC DR PROSITE: PS50862; AA_TRNA_LIGASE_II; 1.
CC KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
  Complete proteome.
CC SQ SEQUENCE 599 AA; 67209 MW; 2B89C952C82A5246 CRC64;
-----
Query Match 33.3%; Score 5; DB 1; Length 599;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 RSTPE 11
ID 11111
DB 174 RSTPE 178
-----
RESULT 130
YEM6_YEAST STANDARD; PRT; 610 AA.
AC P40024;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable ATP-dependent transporter YER036C.
DE YER036C.
GN Saccharomyces cerevisiae (Baker's yeast).
OS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
  Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=S288c / AB972;
RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
  Aviles E., Bero A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
  Chung E., Duncan M., Guzman E., Hartwell G., Hunkeler-Smith S.,
  Hyman R., Kayser A., Kemp C., Lashkari D., Lew H., Lin D.,
  Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,
  Patel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
  Taylor P., Wei Y., Yellon M., Borstein D., Davis R.W.;
RT Submitted (DEC-1994) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. EF3 SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: U18796; AAB64571.1; -.
CC GSD; S0000838; YER036C.
CC DR InterPro: IPR003439; ABC_transportr.
CC DR Pfam: PF00005; ABC_tran; 2.
-----

```

DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE GTP-binding protein cgp-1.
 GN cgp-1.
 OS *Caenorhabditis elegans*.
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 CC Rhabditidae; Meloidae; Peleoderinae; *Caenorhabditis*.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96104562; PubMed=8524289;
 RA Huang Y.-J., Stoffel R., Tobler H., Mueller F.;
 RT "A newly formed telomere in *Ascaris suum* does not exert a telomere
 position effect on a nearby gene.";
 RL Mol. Cell. Biol. 16:130-134(1996).
 CC -1- SIMILARITY: BELONGS TO THE AGP1/GTPBP1 FAMILY OF GTP-BINDING
 CC PROTEINS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL: LA0996; AAA6431.1; -
 DR InterPro: IPR004160; EFTU_Cterm.
 DR InterPro: IPR004161; EFTU_D2.
 DR InterPro: IPR00795; EF_GTPbind.
 DR Pfam: PF00009; GTP_EFTU; 1.
 DR Pfam: PF03143; GTP_EFTU_D3; 1.
 DR Pfam: PF03144; GTP_EFTU_D2; 1.
 KM GTP-binding.
 FT NP_BIND 147 154 GTP (POTENTIAL).
 FT NP_BIND 232 236 GTP (POTENTIAL).
 FT NP_BIND 289 292 GTP (POTENTIAL).
 SO SEQUENCE 593 AA; 65891 MW; B7BB1E586E60CA4 CRC64;

 Query Match 33.3%; Score 5; DB 1; Length 593;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 7 RSTPE 11
 DB 42 RSTPE 46

 RESULT 127
 SYD_MYCTU STANDARD; PRT; 594 AA.
 AC Q50649;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--tRNA ligase)
 DE (AspRS).
 GN ASPs OR RV2572C OR MT2648 OR MTCY227.29.
 OS *Mycobacterium tuberculosis*.
 CC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 CC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RA MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigmler K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,

RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of *Mycobacterium tuberculosis* from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., Deboy R., Dodson R., Gwin M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Decher A., Utterback T., Weidman J., Kouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of *Mycobacterium tuberculosis* clinical and
 RT laboratory strains.";
 RL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) -> AMP +
 CC diphosphate + L-aspartyl-tRNA(Asp).
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL: Z77724; CAB01271.1; ALT_INIT.
 DR EMBL: AE007099; AAK46961.1; -
 DR HSSP: P36419; 1EFW.
 DR TIGR: MT2648; -
 DR TubercuList: RV2572C; -
 DR InterPro: IPR002106; tRNA_ligaseII.
 DR InterPro: IPR004524; ASPs_bact.
 DR InterPro: IPR004115; GAD_dom.
 DR InterPro: IPR004364; tRNA-synt_2.
 DR InterPro: IPR002312; tRNA-synt_2.
 DR InterPro: IPR004365; tRNA-antL.
 DR Pfam: PF00152; tRNA-synt_2; 2.
 DR Pfam: PF01336; tRNA-antL; 1.
 DR Pfam: PF02938; GAD; 1.
 DR PRINTS: PR01042; TRNASYNTHTASP.
 DR TIGRFAMS: TIGR00459; asps_bact.1.
 DR PROSITE: PS50862; AA-TRNA_LIGASE-II; 1.
 KM Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KM Complete proteome.
 SO SEQUENCE 594 AA; 64980 MW; D0013840069FE283 CRC64;

 Query Match 33.3%; Score 5; DB 1; Length 594;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 7 RSTPE 11
 DB 169 RSTPE 173

 RESULT 128
 TREZ_RHISP STANDARD; PRT; 596 AA.
 ID TREZ_RHISP
 AC Q53238;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Maltotriose trehalohydrolase (EC 3.2.1.141) (MTHase) (4-
 DE alpha-D-(1->4)-alpha-D-glucanotrehalohydrolase)
 DE (Maltotriosyl trehalose trehalohydrolase).
 GN TREZ.
 OS *Rhizobium* sp. (strain M-11).
 OS Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--tRNA ligase)
 DE ASPRS (Antigen T5).
 GN ASPS OR M0501 OR MLCB1259.19.
 OS Mycobacterium leprae.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1769;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96071886; PubMed=7591123;
 RA Wieses B., Splierings E., van Noort J., Naafs B., Offringa R.,
 RA Ottenhoff T.;
 RT "Molecular characterization and T-cell-stimulatory capacity of
 RT Mycobacterium leprae antigen T5."
 RL Infect. Immun. 63:4682-4685(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TN;
 RX MEDLINE=21128732; PubMed=11234002;
 RA Cole S.T., Eigmeier K., Parkhill J., James K.D., Thomson N.R.,
 RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
 RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
 RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
 RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
 RA Rutter S., Seeger K., Simon S., Smeadows M., Skellon J., Squares R.,
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
 RA Barrall B.G.;
 RT "Massive gene decay in the leprosy bacillus";
 RL Nature 409:1007-1011(2001).
 CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +
 CC diphosphate + L-aspartyl-tRNA(Asp).
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X77655; CAA54735.1; -;
 DR EMBL: S62268; AAC27132.1; -;
 DR EMBL: AL023591; CAA19094.1; -;
 DR EMBL: AL583918; CAC30009.1; -;
 DR PIR: S42047; S42047.
 DR HSSP: P36419; 1EFW.
 DR Leproma: M0501; -;
 DR InterPro: IPR002106; AAcRNA_ligaseII.
 DR InterPro: IPR004524; AspS_bact.
 DR InterPro: IPR004115; GAD_dom.
 DR InterPro: IPR004364; tRNA-synt_2.
 DR InterPro: IPR002312; tRNA-synt_asp.
 DR InterPro: IPR004365; tRNA_antl.
 DR Pfam: PF00152; tRNA-synt_2; 2.
 DR Pfam: PF01336; tRNA_antl; 1.
 DR Pfam: PF02938; GAD; 1.
 DR PRINTS: PRO1042; TRNASYNTHASP.
 DR TIGRFS: TIGR00459; aspS_bact; 1.
 DR PROSITE: PS50862; AA_TRNA_LIGASE_II; 1.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Complete proteome.
 SQ SEQUENCE 589 AA; 64630 MW; 6B0EA0675367031C CRC64;

DB 166 RSTPE 170
 II III
 RESULT 125
 FTSZ_BARBA STANDARD; PRT; 592 AA.
 ID FTSZ_BARBA
 AC 031314;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cell division protein ftsz (75 kDa antigen).
 GN FTSZ.
 OS Bartonella bacilliformis.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Bartonellaceae; Bartonella.
 OX NCBI_TaxID=774;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KC584;
 RX MEDLINE=97369823; PubMed=9226264;
 RA Padmalayam I., Anderson B., Kron M., Kelly T., Baumstark B.;
 RT "The 75-kilodalton antigen of Bartonella bacilliformis is a
 RT structural homolog of the cell division protein ftsz.";
 RL J. Bacteriol. 179:4545-4552(1997).
 RN [2]
 RP REVISIONS.
 RA Kelly T.;
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: This protein is essential to the cell-division process.
 CC it seems to assemble into a dynamic ring on the inner surface of
 CC the cytoplasmic membrane at the place where division will occur,
 CC and the formation of the ring is the signal for septation to
 CC begin. Binds to and hydrolyzes GTP (by similarity).
 CC -1- SUBUNIT: Aggregates to form a ring-like structure (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic. Assembles at the inner surface
 CC of the cytoplasmic membrane (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF007266; AAC15082.1; -;
 DR HSSP: Q57816; 1FSZ.
 DR InterPro: IPR000158; FtsZ.
 DR InterPro: IPR003008; Tubulin_FtsZ.
 DR Pfam: PF00091; tubulin; 1.
 DR PRINTS: PRO0423; CELDIVISFTSZ.
 DR TIGRFS: TIGR00065; ftsz; 1.
 DR PROSITE: PS01134; FTSZ_1; 1.
 DR PROSITE: PS01135; FTSZ_2; 1.
 KW Cell division; Septation; GTP-binding.
 FT NP_BIND 107 115 GTP (POTENTIAL).
 SQ SEQUENCE 592 AA; 63503 MW; C32007DADCD2D75B CRC64;

Query Match 33.3%; Score 5; DB 1; Length 592;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PHRST 9
 II III
 DB 557 PHRST 561
 RESULT 126
 CGPI_CAEEL STANDARD; PRT; 593 AA.
 ID CGPI_CAEEL
 AC 018905;

DR PROSITE: PS00031; NUCLEAR_RECEPTOR; FALSE_NEG.
KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
FT ZINC_FINGER 96 166 NUCLEAR_RECEPTOR-TYPE.
FT ZN_FING 96 117 C4-TYPE.
FT ZN_FING 133 161 C4-TYPE.
SQ SEQUENCE 579 AA; 66071 MW; C76F1646819ED4CC CRC64;

Query Match 33.3%; Score 5; DB 1; Length 579;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 6 HRSTP 10
|||||
DB 253 HRSTP 257

RESULT 122

SYD_THEME STANDARD; PRT; 579 AA.
ID SYD_THEME
AC O9X1F4;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--tRNA ligase)
DE (AspRS).
GN ASPS OR TM1441.
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogae (class); Thermotogales;
OC Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=MSB / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,
Halt D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
McDonald L., Utterback T.D., Malek J.A., Linher K.D., Garrett M.M.,
Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
genome sequence of Thermotoga maritima."
RL Nature 399:323-329 (1999).
CC -1 CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +
diphosphate + L-aspartyl-tRNA(Asp).
CC -1 SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1 SUBCELLULAR LOCATION: Cytoplasmic.
CC -1 SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001796; AAD36510.1; -
DR HSSP: P36419; LEFW.
DR TIGR: TM1441.
DR InterPro: IPR002106; AATRNA_ligaseII.
DR InterPro: IPR004524; ASPS_bact.
DR InterPro: IPR004115; GAD_dom.
DR InterPro: IPR004364; tRNA-synt_2.
DR InterPro: IPR002312; tRNA-synt_asp.
DR InterPro: IPR004365; tRNA_antl.
DR Pfam: PF00152; tRNA-synt_2; 2.
DR Pfam: PF01336; tRNA_antl; 1.
DR Pfam: PF02938; GAD; 1.
DR PRINTS: PRO1042; TRNASYNTASP.
DR TIGRfams: TIGR00459; asps_bact; 1.
DR PROSITE: PS50862; AA_TRNA_LIGASE_II; 1.

KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
SQ SEQUENCE 579 AA; 66495 MW; 841951477100B9A61 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 579;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 7 RSTPE 11
|||||
DB 167 RSTPE 171

RESULT 123

KM8S_YEAST STANDARD; PRT; 586 AA.
ID KM8S_YEAST
AC Q03533;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable serine/threonine-protein kinase YMR291W (EC 2.7.1.-).
GN YMR291W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972.
RA Badcock K., Churher C., Barrell B.G., Rajandream M.A.;
RL Submitted (Aug-1994) to the EMBL/GenBank/DDBJ databases.
CC -1 SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X80836; CA556800.1; -
DR SGD: S0004905; YMR291W.
DR InterPro: IPR00719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW Hypothetical protein; Transferase; Serine/threonine-protein kinase;
FT ATP-binding. 39 351 PROTEIN_KINASE.
FT DOMAIN 39 351
FT NP_BIND 45 53 ATP (BY SIMILARITY).
FT BINDING 68 68 ATP (BY SIMILARITY).
FT ACT_SITE 180 180 BY SIMILARITY.
SQ SEQUENCE 586 AA; 66219 MW; 584D620D5F798EF6 CRC64;
Query Match 33.3%; Score 5; DB 1; Length 586;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 9 TPESR 13
|||||
DB 504 TPESR 508

RESULT 124

SYD_MYCLE STANDARD; PRT; 589 AA.
ID SYD_MYCLE
AC P36429; P95671;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)

DB 227 RSTPE 231

RESULT 120

GPV_RAT STANDARD; PRT; 567 AA.

AC 008770;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Platelet glycoprotein V precursor (GPV) (CD42D).

GN GP5.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Wistar; TISSUE=Liver;

RX MEDLINE=97275136; PubMed=9129030;

RA Ravant C., Morales M., Azores D.O., Moog S., Schubler S., Grunert P., Loew D., van Dorsselaer A., Gazenave J.-P., Lanza F.;

RT "Gene cloning of rat and mouse platelet glycoprotein V: identification of megakaryocyte-specific promoters and demonstration of functional thrombin cleavage.";

RL Blood 89:3253-3262(1997).

CC -1- FUNCTION: THE GPIIb-IIIa COMPLEX FUNCTIONS AS THE VON WILLEBRAND FACTOR RECEPTOR AND MEDIATES VON WILLEBRAND FACTOR-DEPENDENT PLATELET ADHESION TO BLOOD VESSELS. THE ADHESION OF PLATELETS TO INJURED VASCULAR SURFACES IN THE ARTERIAL CIRCULATION IS A CRITICAL INITIATING EVENT IN HEMOSTASIS (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: TYPE I membrane protein.

CC -1- SIMILARITY: CONTAINS 14 LEUCINE-RICH REPEATS (LRR).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; Z69594; CAA3440.1; -

DR InterPro: IPR001611; LRR.

DR InterPro: IPR000483; LRR_Cterm.

DR InterPro: IPR000372; LRR_Nterm.

DR InterPro: IPR003582; LRR_out.

DR InterPro: IPR003591; LRR_typ.

DR Pfam: PF00560; LRR_13.

DR Pfam: PF01463; LRRCT; 1.

DR PRINTS: PRO0019; LEURICHRPT.

DR SMART: SM00370; LRR; 2.

DR SMART: SM00082; LRRCT; 1.

DR SMART: SM00013; LRRNT; 1.

DR SMART: SM00369; LRR_typ; 10.

KW Platelet; Transmembrane; Glycoprotein; Blood coagulation; Repeat; Leucine-rich repeat; Cell adhesion; Signal.

FT SIGNAL 1 16

FT CHAIN 17 567

FT DOMAIN 17 522

FT TRANSMEM 523 543

FT DOMAIN 544 567

FT REPEAT 73 96

FT REPEAT 97 120

FT REPEAT 122 144

FT REPEAT 145 168

FT REPEAT 169 192

FT REPEAT 194 216

FT REPEAT 217 240

FT REPEAT 241 264

FT REPEAT 266 288

FT REPEAT 289 312

FT REPEAT 314 337

LRR 1.

LRR 2.

LRR 3.

LRR 4.

LRR 5.

LRR 6.

LRR 7.

LRR 8.

LRR 9.

LRR 10.

LRR 11.

FT REPEAT 338 361

FT REPEAT 362 385

FT REPEAT 387 409

FT CARBOHYD 51 51

FT CARBOHYD 181 181

FT CARBOHYD 243 243

FT CARBOHYD 298 298

FT CARBOHYD 312 312

FT CARBOHYD 385 385

FT CARBOHYD 498 498

SO SEQUENCE 567 AA; 63344 MW; CA10708E0D03707F CRC64;

Query Match 33.3%; Score 5; DB 1; Length 567;

Best Local Similarity 100.0%; Pred. No. 2.2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRA 14

DB 454 PESRA 458

RESULT 121

NH22_CAEEL STANDARD; PRT; 579 AA.

ID NH22_CAEEL

AC 009587; O9GFI4; O9GTI5;

DT 01-FEB-1996 (Rel. 33, Created)

DT 15-FEB-1998 (Rel. 37, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Nuclear hormone receptor family member nhr-22.

GN NHR-22 OR K06A1.4.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RA Fulton L.;

RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.

RN [2]

RP REVISIONS.

RA Waterston R.;

RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE OF 25-579 FROM N.A.

RA Bogan A., Maine C.V., Yamamoto K., Cohen F., Sluder A.E.;

RT "Caenorhabditis elegans nuclear receptor sequences exhibit biophysical compatibility with the ligand-binding domain fold.";

RT Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: ORPHAN NUCLEAR RECEPTOR.

CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).

CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; U23449; AAC34301.1; -

DR EMBL; AF273772; AAG15121.1; -

DR EMBL; AF273773; AAG15122.1; -

DR HSSP; P20393; 1A6Y.

DR WormPeP; K06A1.4; CE18017.

DR InterPro: IPR000536; Hormone_rec_lig.

DR InterPro: IPR001628; Znf_C4steroid.

DR Pfam; PF00104; hormone_rec; 3.

DR Pfam; PF00105; zif-C4; 3.

DR ProDom; PD0000035; Znf_C4steroid; 1.

DR SMART; SM00430; HOLI; 1.

DR SMART; SM00399; Znf_C4; 1.

Best Local Similarity 100.0%; Pred. NO. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
|||||
DB 160 RSTPE 164

RESULT 118

HEMA_IJAP STANDARD: PRT: 562 AA.

AC P03431.
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hemagglutinin precursor [Contains: Hemagglutinin HAI chain;
Hemagglutinin HAZ chain].

OS Influenza A virus (strain A/Japan/305/57').
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.

OX NCBI_TaxID=11421;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=81030852; PubMed=7421990;
RA Getting M.-J., Bye J., Skehel J.J., Waterfield M.;

RT "Cloning and DNA sequence of double-stranded copies of haemagglutinin
genes from H2 and H3 strains elucidates antigenic shift and drift in
human influenza virus".

RL Nature 287:301-306(1980).

CC -1- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CELL RECEPTORS AND FOR INITIATING INFECTION.

CC -1- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
(HA1 AND HA2) LINKED BY A DISULFIDE BOND.

CC -1- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>
or send an email to license@sdb-sdb.ch).

CC EMBL: J02127; AAA43185.1; -.

DR PIR: A04062; HMIY2.

DR HSSP: P03437; IHTM.

DR InterPro: IPR001364; Hemagglut.

DR Pfam: PF00509; Hemagglutinin; 1.

DR PRINTS: PR00329; HEMAGGLUTN12.

DR PRODOM: PD000225; Hemagglutn; 1.

KW Envelope protein; Hemagglutinin; glycoprotein; Signal.

FT SIGNAL 1 15

FT CHAIN 16 339 HEMAGGLUTININ HAI CHAIN.

FT CHAIN 341 562 HEMAGGLUTININ HAZ CHAIN.

FT CAROHD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CAROHD 26 26 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CAROHD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CAROHD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CAROHD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CAROHD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CAROHD 494 494 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CAROHD 553 553 N-LINKED (GLCNAC. . .) (POTENTIAL).

SO SEQUENCE 562 AA; 63118 MW; 687FDC0389931630 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 562;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
|||||
DB 222 RSTPE 226

RESULT 119

SIS2_YEAST STANDARD: PRT: 562 AA.

AC P36024.
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE SIS2 protein (Halotolerance protein HAL3).

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

OX NCBI_TaxID=4932;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=95220693; PubMed=7705654;

RA di Como C.J., Bose R., Arndt K.T.;

RT "Overexpression of SIS2, which contains an extremely acidic region,
increases the expression of SWI4, CLN1 and CLN2 in *slt4* mutants.";

RL Genetics 139:95-107(1995).

RN [2]

RP SEQUENCE FROM N.A.

RA Pohl T.M., Pohl F.M.;

RT Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.

RN [3]

RP MEDLINE=96009574; PubMed=7565698;

RA Ferrando A., Kton S.J., Rios G., Pink G.R., Serrano R.;

RT "Regulation of cation transport in *Saccharomyces cerevisiae* by the
salt tolerance gene HAL3.";

RL Mol. Cell. Biol. 15:5470-5481(1995).

RN [4]

RP IDENTIFICATION AS INHIBITORY SUBUNIT OF PP2L.

RX MEDLINE=96301578; PubMed=9636153;

RA de Nadal E., Clotet J., Posas F., Gomez N., Arino J.;

RT "The yeast halotolerance determinant Hal3p is an inhibitory subunit of
the Ppz1 Ser/Thr protein phosphatase.";

RL Proc. Natl. Acad. Sci. U.S.A. 95:7357-7362(1998).

CC -1- FUNCTION: MAY STIMULATE EXPRESSION OF CERTAIN GENES THAT ARE
PERIODICALLY EXPRESSED DURING LATE G1. ALSO MODULATES THE

EXPRESSION OF THE ENAI ATPASE. INTERACTS WITH THE C-TERMINAL
DOMAIN OF THE SERINE-THREONINE PROTEIN PHOSPHATASE PP2L AND ACTS

AS AN INHIBITORY SUBUNIT OF PP2L.

CC -1- SUBCELLULAR LOCATION: NUCLEAR OR CYTOPLASMIC.

CC -1- SIMILARITY: TO C-TROPICALIS SIS2/HAL3 AND TO YEAST YKL088W. SOME,
TO A. THALIANA HAL3A AND HAL3B.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>
or send an email to license@sdb-sdb.ch).

CC EMBL: U01878; AAA80000.1; -.

DR EMBL: Z28297; CAAB2151.1; -.

DR PIR: S38149; S38149.

DR HSSP: O9SWE5; 1R30.

DR SGD: S0001780; SIS2.

DR InterPro: IPR003382; Flavoprotein.

DR Pfam: PF02441; Flavoprotein; 1.

KW Protein phosphatase inhibitor; Nuclear protein.

FT DOMAIN 496 553 ASP/GLU-RICH (HIGHLY ACIDIC).

SO SEQUENCE 562 AA; 62478 MW; 19A9A475145DA7AB CRC64;

Query Match 33.3%; Score 5; DB 1; Length 562;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
|||||

DR EMBL: L34837; AAA53129.1; -
DR EMBL: Z54141; CA90837.1; -
DR SGD: S0004938; FET4. -
KW Transmembrane; Transport; Iron transport.
FT TRANSMEM 98 118 POTENTIAL.
FT TRANSMEM 226 246 POTENTIAL.
FT TRANSMEM 272 292 POTENTIAL.
FT TRANSMEM 355 375 POTENTIAL.
FT TRANSMEM 384 404 POTENTIAL.
FT TRANSMEM 466 486 POTENTIAL.
FT TRANSMEM 494 514 POTENTIAL.
FT TRANSMEM 283 283 V -> I (IN REF. 1).
FT CONFLICT 441 441 F -> L (IN REF. 1).
FT CONFLICT 450 450 T -> I (IN REF. 1).
SQ SEQUENCE 552 AA; 62792 MW; 9AB1AE5DC6BA08E CRC64;

Query Match 33.3%; Score 5; DB 1; Length 552;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPE 11
Db 173 RSTPE 177

RESULT 116
RORG_HUMAN STANDARD; PRT; 560 AA.
ID RORG_HUMAN
AC P51449;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Nuclear receptor ROR-gamma (Nuclear receptor RZR-gamma).
GN RORC OR NR1P3 OR RORC OR RZRG.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Skeletal muscle;
RX MEDLINE=9510350; PubMed=7811290;
RA Hirose T., Smith R.J., Jetten A.M.;
RT "ROR gamma: the third member of ROR/RZR orphan receptor subfamily
RT that is highly expressed in skeletal muscle."
RL Biochem. Biophys. Res. Commun. 205:1976-1983(1994).
CC -1- FUNCTION: ORPHAN NUCLEAR RECEPTORS.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN SKELETAL MUSCLE.
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
CC NRI SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: U16997; AAA64751.1; -
DR HSSP: P20393; 1A6Y.
DR TRANSFAC: T02749; -
DR Genew: HGNC:10260; RORC.
DR MIM: 602943; -
DR InterPro: IPR000536; Hormone_rec_11g.
DR InterPro: IPR001628; znf_C4steroid.
DR Pfam: PF00104; hormone_rec_1.
DR Pfam: PF00105; znf_C4_1.
DR PRINTS: PR00047; STROIDFINGER.
DR PRODom: PD000035; znf_C4steroid_1.
DR SMART: SM00430; HOL1.1.
DR SMART: SM00399; znf_C4_1.
DR

DR PROSITE: PS00031; NUCLEAR_RECEPTOR_1.
KW Transcription regulation; DNA-binding; Nuclear protein;
KW Zinc-finger.
FT DOMAIN 1 30 MODULATING (POTENTIAL).
FT DNA_BIND 31 96 NUCLEAR RECEPTOR-TYPE.
FT ZN_FING 31 51 C4-TYPE.
FT ZN_FING 67 91 C4-TYPE.
FT DOMAIN 97 268 HINGE (POTENTIAL).
FT DOMAIN 269 560 LIGAND-BINDING (POTENTIAL).
FT DOMAIN 121 130 POLY-GLN.
SQ SEQUENCE 560 AA; 62589 MW; 25D05C0AC83C3736 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 560;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPE 11
Db 257 RSTPE 261

RESULT 117
TARA_MOUSE STANDARD; PRT; 561 AA.
ID TARA_MOUSE
AC O99KW3;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Protein Tara (Trio-associated repeat on actin) (Fragment).
GN TARA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Mammary gland;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DDA databases.
CC -1- FUNCTION: May regulate actin cytoskeletal organization, cell
CC spreading and cell contraction by directly binding and stabilizing
CC filamentous (F)-actin. The localized formation of Tara and Trio
CC complexes coordinates the amount of F-actin present in stress
CC fibers. May also serve as a linker protein to recruit proteins
CC required for F-actin formation and turnover.
CC -1- SUBUNIT: Binds to Trio and F-actin. May also interact with myosin
CC II (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic; localized to F-actin in a
CC periodic pattern (By similarity).
CC -1- DOMAIN: Contains at least 2 actin-binding sites per coiled-coil
CC dimer (By similarity).
CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: BC003984; AAH03984.1; -
DR InterPro: IPR001849; PH.
DR Pfam: PF00169; PH; 1.
DR SMART: SM00233; PH; 1.
DR PROSITE: PS50003; PH_DOMAIN; 1.
KW Cytoskeleton; Actin-binding; Coiled coil.
FT NON_TER 1 1
FT DOMAIN <1 83 PH.
FT DOMAIN 258 560 COILED COIL (POTENTIAL).
SQ SEQUENCE 561 AA; 64762 MW; 436DFA47B3AC25 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 561;

DR InterPro: IPR000822; Znf_C2H2.
DR Pfam: PF00096; 2f_C2H2; 7.
DR SMART; SM00349; KRAB; 1.
DR SMART; SM00355; znf_C2H2; 7.
DR PROSITE; PS00805; KRAB; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 7.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 7.
DR Transcription regulation; DNA-binding; zinc-finger; Metal-binding;
KW Nuclear protein; Repeat.
FT DOMAIN 1 64 KRAB.
FT DOMAIN 225 457 ZINC FINGERS.
FT ZN_FING 225 247 C2H2-TYPE.
FT ZN_FING 253 275 C2H2-TYPE.
FT ZN_FING 281 303 C2H2-TYPE.
FT ZN_FING 309 331 C2H2-TYPE.
FT ZN_FING 337 359 C2H2-TYPE.
FT ZN_FING 365 387 C2H2-TYPE.
FT ZN_FING 435 457 C2H2-TYPE.
SQ SEQUENCE 543 AA; 61772 MW; ADD987504ECAC019 CRC64;

Query Match
Best Local Similarity 33.3%; Score 5; DB 1; Length 543;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRA 14
Db 501 PESRA 505

RESULT 114
MERA_PSEFL STANDARD; PRT; 548 AA.
AC 051772;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mercuric reductase (EC 1.16.1.1) (Hg(II) reductase).
GN MERA.
OS Pseudomonas fluorescens.
OC Plasmid pMER327.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=294;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94341572; PubMed=8063107;
RA Hobman J., Kholodil G., Nikiforov V., Ritchie D.A., Strike P.,
RA Yurleva O.;
RA "The sequence of the mer operon of pMER327/419 and transposon ends of
RT pMER327/419, 330 and 05.";
RL Gene 146:73-78(1994)
CC -1- FUNCTION: RESISTANCE TO HG(2+) IN BACTERIA APPEARS TO BE GOVERNED
CC BY A SPECIALIZED SYSTEM WHICH INCLUDES MERCURIC REDUCTASE. MERA
CC PROTEIN IS RESPONSIBLE FOR VOLATILIZING MERCURY AS HG(0).
CC -1- CATALYTIC ACTIVITY: Hg + NADP(+) + H(+) = Hg(2+) + NADPH.
CC -1- COFACTOR: FAD.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- MISCELLANEOUS: THE ACTIVE SITE IS A REDOX-ACTIVE DISULFIDE BOND.
CC -1- SIMILARITY: BELONGS TO THE PYRIDINE NUCLEOTIDE-DISULFIDE
CC OXIDOREDUCTASES CLASS-I.
CC -1- SIMILARITY: CONTAINS 1 HMA DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X73112; CAA51542.1; -

DR HSP, P04129; 1AFJ.
DR InterPro: IPR001327; FAD_Pyr_redux.
DR InterPro: IPR001934; HeavyMe_transpl.
DR InterPro: IPR000815; Hg_reductase.
DR InterPro: IPR001100; Pyr_redux.
DR InterPro: IPR004099; Pyr_redux_dim.
DR Pfam: PF00070; Pyr_redux; 1.
DR Pfam: PF00403; HMA; 1.
DR Pfam: PF02852; Pyr_redux_dim; 1.
DR PRINTS; PR00368; FADPNR.
DR PRINTS; PR00945; HGRDPTASE.
DR PRINTS; PR00411; PNDRTASEI.
DR ProDom; PD000139; FAD_Pyr_redux; 1.
DR PROSITE; PS01047; HMA_1; 1.
DR PROSITE; PS00846; HMA_2; 1.
DR PROSITE; PS00076; PYRIDINE_REDOX_1; 1.
KW Mercuric resistance; Oxidoreductase; Flavoprotein; FAD; NADP;
KW Mercury; Redox-active center; Metal-binding; Plasmid.
FT DOMAIN 1 66 HMA.
FT NP_BIND 87 117 FAD (ADP PART) (PROBABLE).
FT DISULFID 123 128 REDOX-ACTIVE.
FT NP_BIND 380 390 FAD (FLAVIN PART) (BY SIMILARITY).
FT METAL 545 545 HG(2+) (POTENTIAL).
FT METAL 546 546 HG(2+) (POTENTIAL).
SQ SEQUENCE 548 AA; 57566 MW; C3AC40203F9E86A7 CRC64;

Query Match
Best Local Similarity 100.0%; Score 5; DB 1; Length 548;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRA 14
Db 36 PESRA 40

RESULT 115
FET4_YEAST STANDARD; PRT; 552 AA.
AC P40988;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Low-affinity Fe(II) transport protein.
GN FET4 OR YMR319C OR YMR924.11C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95014434; PubMed=7929320;
RA Dix D.R., Bridgman J.T., Broderius M.A., Byersdorfer C.A.,
RA Eide D.J.;
RA "The FET4 gene encodes the low affinity Fe(II) transport protein of
RT Saccharomyces cerevisiae.";
RL J. Biol. Chem. 269:26092-26099(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Churcher C.M., Louis E.J., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: REQUIRED FOR FERROUS IRON LOW AFFINITY UPTAKE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- INDUCTION: BY IRON DEPRIVATION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC

```

DR Pfam: PF00319; SRF-TF; 1.
DR PRINTS: PR00404; MADS-DOMAIN.
DR SMART: SM00432; MADS: 1.
DR PROSITE: PS00350; MADS_BOX_1: 1.
DR PROSITE: PS00066; MADS_BOX_2: 1.
KW Transcription regulation; Nuclear protein; DNA-binding; Activator;
KW Multigene family.
FT DOMAIN 3 57 MADS.
FT DNA BIND 58 86 MER2-TYPE (POTENTIAL).
FT DOMAIN 3 31 ARG/LYS-RICH (BASIS).
FT DOMAIN 182 225 SER/THR-RICH.
FT DOMAIN 451 460 GLN/HIS-RICH.
FT DOMAIN 466 477 SER-RICH.
SQ SEQUENCE 516 AA; 56403 MW; 8C85505AC123F4E CRC64;

Query Match
Best Local Similarity 33.3%; Score 5; DB 1; Length 516;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGP 5
DB 277 SHLGP 281

RESULT 112
NAB2 MOUSE STANDARD; PRT; 525 AA.
AC 061127;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE NGFI-A binding protein 2 (EGR-1 binding protein 2).
GN NAB2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Brain;
RX MEDLINE=96251303; PubMed=8668170;
RA Svaren J., Sevelson B.R., Apel E.D., Zimonjic D.B., Popescu N.C.,
RA Milbrandt J.;
RA "NAB2, a corepressor of NGFI-A (Egr-1) and Krox20, is induced by
RT proliferative and differentiative stimuli.";
RT Mol. Cell. Biol. 16:3545-3553(1996).
RN [2]
RP PARTIAL SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=97271553; PubMed=916479;
RA Svaren J., Apel E.D., Simburger K.S., Jenkins N.A., Gilbert D.J.,
RA Copeland N.A., Milbrandt J.;
RA "The NAB2 and stat6 genes share a common transcription termination
RT region.";
RT Genomics 41:33-39(1997).
CC -1- FUNCTION: ACTS AS A TRANSCRIPTIONAL REPRESSOR FOR ZINC FINGER
CC TRANSCRIPTION FACTORS EGR1 AND EGR2. ISOFORM 2 LACKS REPRESSION
CC ABILITY.
CC -1- SUBUNIT: HOMOMULTIMERIS MAY ASSOCIATE WITH EGR1 BOUND TO DNA (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR. ISOFORM 2 IS NOT LOCALIZED TO
CC THE NUCLEUS.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: 1 (SHOWN HERE) AND 2; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN BRAIN AND THYMUS, AND AT
CC LOWER LEVELS IN SPLEEN, KIDNEY, HEART AND TESTIS. ISOFORM 1 IS
CC PREDOMINANTLY EXPRESSED IN TESTIS, WHEREAS ISOFORM 3 IS MORE
CC ABUNDANT IN THYMUS.
CC -1- INDUCTION: BY SERUM STIMULATION.
CC -1- DOMAIN: THE NAB CONSERVED DOMAIN 1 (NCDD1) INTERACTS WITH EGR1
CC INHIBITORY DOMAIN AND MEDIATES MULTIMERIZATION.
CC -1- DOMAIN: THE NAB CONSERVED DOMAIN 2 (NCDD2) IS NECESSARY FOR
CC TRANSCRIPTIONAL REPRESSION.
CC -1- SIMILARITY: BELONGS TO THE NAB FAMILY.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U47543; AAC52650.1; -.
DR MGD; MG1:107563; Nab2.
KW Transcription regulation; Repressor; Alternative splicing.
FT DOMAIN 35 113 NCDD1.
FT DOMAIN 267 356 NCDD2.
FT DOMAIN 353 384 NECESSARY FOR NUCLEAR LOCALIZATION (BY
FT VARSPLIC 320 322 SIMILARITY).
FT VARSPLIC 323 525 LFT -> ASL (IN ISOFORM 2).
SQ SEQUENCE 525 AA; 56653 MW; FED42BE94A8BD804 CRC64;

Query Match
Best Local Similarity 33.3%; Score 5; DB 1; Length 525;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STRES 12
DB 189 STRES 193

RESULT 113
ZN08 HUMAN STANDARD; PRT; 543 AA.
AC P17098;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Zinc finger protein 8 (Zinc finger protein HF.18) (Fragment).
GN ZNF8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=90169993; PubMed=2106481;
RA Lania L., Donli E., Pannuti A., Pascucci A., Pengue G.,
RA Feliciello I., la Mantia G., Ianfrancone L., Pelicci P.-G.;
RT "cDNA isolation, expression analysis, and chromosomal localization of
RT two human zinc finger genes.";
RT Genomics 6:333-340(1990).
CC -1- FUNCTION: MAY FUNCTION AS A TRANSCRIPTION FACTOR.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- TISSUE SPECIFICITY: UBICITOUSLY PRESENT IN MANY HUMAN CELL LINES
CC OF DIFFERENT EMBRYOLOGICAL DERIVATION.
CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -1- SIMILARITY: CONTAINS 1 KRAE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M29581; AAA61314.1; -.
DR PIR: B34612; B34612.
DR HSSP: P08046; 1A1H.
DR Genew; HGNC:13134; ZNF8.
DR MIM: 194532; -.
DR Interpro; IPR001909; KRAE.

```

```

RA Perret X.;
RT "Molecular basis of symbiosis between Rhizobium and legumes.";
RL Nature 387:394-401(1997).
CC -1- SIMILARITY: TO A NUMBER OF ORF OF VARIABLE SIZE IN R.MELILOTI AND
CC A.TUMEFACIENS TI PLASMID.
CC -1- SIMILARITY: TO YAAO, Y4HP AND Y4OI.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AE000079; AAB91716.1; -
CC InterPro: IPR004291; Transposase_25.
CC Pfam: PF03050; Transposase_25; 1.
CC Hypothetical protein; Plasmid.
CC SEQUENCE 511 AA; 56701 MW; BB67FA8B9DE9601 CRC64;
SQ
Query Match 33.3%; Score 5; DB 1; Length 511;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 10 PESRA 14
DB 101 PESRA 105
RESULT 110
YAT3_SCHPO STANDARD: PRT; 513 AA.
ID YAT3_SCHPO
AC Q10150;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein C1D4.03c in chromosome I.
GN SPAC1D4.03c.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21848401; PubMed=11859360;
RX Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moyle S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robden J., Grynoprez B.,
RA Wajsbom I., Vansteleels E., Rieger M., Schaefer M., Meller-Auer S.,
RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purrelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Leclaire V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucchi M., Rochet M., Galliard C., Tallada V.A., Garçon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
```

```

RL Nature 415:871-880(2002).
CC -1- SIMILARITY: TO YEAST YGL124C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: Z69239; CA93212.1; -
CC InterPro: IPR004353; Yeast73DUF.
CC Pfam: PF03164; DUF254; 1.
CC PRINTS: PR01546; YEAST73DUF.
CC Hypothetical protein.
CC SEQUENCE 513 AA; 58003 MW; 17458C42CD5BB10A CRC64;
SQ
Query Match 33.3%; Score 5; DB 1; Length 513;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 RSTPE 11
DB 53 RSTPE 57
RESULT 111
MEFA_XENLA STANDARD: PRT; 516 AA.
ID MEFA_XENLA
AC Q03414;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Myocyte-specific enhancer factor 2A homolog (Serum response factor-
DE like protein 2) (SL-2).
GN MEFA2 OR SL2.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Neurula;
RX MEDLINE=93099873; PubMed=1281451;
RA Chambers A.E., Kotecha S., Towers N., Mohun T.J.;
RT "Muscle-specific expression of SRF-related genes in the early embryo
RT of Xenopus laevis."
RL EMBO J. 11:4981-4991(1992).
CC -1- FUNCTION: MAY REGULATE MUSCLE-SPECIFIC TRANSCRIPTION IN THE
CC EMBRYO AND MAY REGULATE TRANSCRIPTION OF A VARIETY OF CELL TYPES
CC IN THE ADULT. IT BINDS TO THE SEQUENCE CTA/T/A/ATAR.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: RESTRICTED TO THE SOMITIC MESODERM OF EARLY
CC EMBRYOS AND TO THE BODY MUSCLE (MYOTOMES) OF THE TADPOLE.
CC EXPRESSED IN ALL TISSUES EXAMINED IN THE ADULT.
CC -1- DEVELOPMENTAL STAGE: EXPRESSION BEGINS IN THE EARLY NEURULA.
CC -1- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC FACTORS. MEF2 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: Z19123; CA979530.1; -
CC HSSP: P11831; 1SRS.
CC TRANSFAC: T01784; -
CC InterPro: IPR002100; TF_MADSbox.
```

RA Pacinani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hohnselt J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Machado M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Pelxoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silva J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen *Xylella fastidiosa*.";
RL Nature 406:151-159(2000)
CC -1- CATALYTIC ACTIVITY: ATP + L-lysine + tRNA(Lys) = AMP + diphosphate
CC + L-lysyl-tRNA(Lys).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AE003947; AAF83922.1; -
DR HSSP: P14825; IE10.
DR InterPro: IPR002106; AATRNA_LigaseI.
DR InterPro: IPR004364; tRNA-synt_2.
DR InterPro: IPR002313; tRNA-synt_2.
DR InterPro: IPR004365; tRNA-anti.
DR Pfam: PF00152; tRNA-synt_2; 1.
DR Pfam: PF01336; tRNA-anti; 1.
DR PRINTS: PR00982; TRNASTNTLYS.
DR TIGRfams: TIGR00499; lysS_bact; 1.
DR PROSITE: PS50862; AA-TRNA_LIGASE_I; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
SQ SEQUENCE 506 AA; 57531 MW; B194E13ED1CC6D0 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 506;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESR 13
|||||
DB 178 TPESR 182

RESULT 108
ID MORE BORBU STANDARD; PRT; 508 AA.
AC OS1219;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE UDP-N-acetylmuramoyl-alanyl-D-glutamate--2,6-diaminopimelate ligase
DE (EC 6.3.2.13) (UDP-N-acetylmuramyl-tripeptide synthetase) (Meso-
DE diaminopimelate-adding enzyme) (UDP-MurNAc-tripeptide synthetase).
GN MORE OR BB0201.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.

OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC SRRATN-ARCC 35210 / B31;
RX MEDLINE=98065943; Pubmed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.F., Fleischmann R.D., Richardson D.,
RA Peterson J., Karpavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Uitterback T., Watney L., McDonald L., Artlich P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RT Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*.";
RL Nature 390:580-586(1997).
CC -1- FUNCTION: CELL WALL FORMATION. DIAMINOPIMELIC ACID ADDING ENZYME
CC (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl-L-alanyl-D-
CC glutamate + meso-2,6-diaminopimelate = ADP + phosphate +
CC UDP-N-acetylmuramoyl-L-alanyl-D-glutamyl-meso-2,6-
CC diaminopimelate.
CC -1- PATHWAY: Peptidoglycan biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE MURDEF FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AE001130; AAC66588.1; -
DR TIGR: BB0201;
DR InterPro: IPR000713; Mur_Ligase.
DR InterPro: IPR004101; Mur_Ligase_C.
DR Pfam: PF01225; Mur_Ligase; 1.
DR Pfam: PF02875; Mur_Ligase_C; 1.
DR TIGRfams: TIGR01085; murE; 1.
KW Peptidoglycan synthesis; Cell wall; Cell division; Ligase;
KW ATP-binding; Complete proteome.
FT NP_BIND 118 124
SQ SEQUENCE 508 AA; 57148 MW; 44043B05F3683BC CRC64;

Query Match 33.3%; Score 5; DB 1; Length 508;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STES 12
|||||
DB 163 STES 167

RESULT 109
ID Y4UD_RHISN STANDARD; PRT; 511 AA.
AC P55504;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 56.7 kDa protein Y4UD.
GN Y4UD.
OS Rhizobium sp. (strain NGR234).
OC Plasmid sym pNGR234a.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=394;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97305956; Pubmed=9163424;
RX Frelberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,

```

CC AN EARLY ROLE IN SPECIFIC ECTODERMAL CELLS, AND A SUBSEQUENT ROLE
CC IN THE EMBRYONIC NERVOUS SYSTEM.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: NEUROBLASTS AND SENSORY ELEMENTS.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED PRIMARILY DURING THE FIRST HALF OF
CC EMBRYOGENESIS. INITIAL EXPRESSION IN CELLULAR BLASTODERM STAGE,
CC THEN IN ECTODERMAL STRIPES DURING GERMBAND EXTENSION. BROAD
CC EXPRESSION IN THE NEUROECTODERM FOLLOWED BY LIMITATION TO DISCRETE
CC SUBSETS OF CNS CELLS. AND EXPRESSION IN SPECIFIC PNS NEURONS AND
CC SUPPORT CELLS.
CC -1- SIMILARITY: BELONGS TO THE POU TRANSCRIPTION FACTOR FAMILY.
CC CLASS-2 SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 1 HOMEBOX DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-stb.ch/announce/
CC or send an email to license@isb-stb.ch).
CC -----
CC EMBL: S80559; AAB21408.1; -
CC EMBL: M65016; AAA28481.1; -
CC EMBL: M81958; AAA28830.2; -
CC EMBL: M93149; AAA28732.1; -
CC EMBL: AE003637; AAF53209.1; -
CC PIR: B41277; B41277.
CC HSP: P14859; 10CT.
CC TRANSFAC: T01901; -
CC FLYBASE: FBgn0004394; pdm2.
CC InterPro: IPR001356; Homeobox.
CC InterPro: IPR000327; POU_domain.
CC Pfam: PF00046; homeobox; 1.
CC PRINTS: PR00157; pou; 1.
CC PRODOM: PD00028; POU_DOMAIN.
CC PRODOM: PD000583; Homeobox; 1.
CC SMART: SM00389; HOX; 1.
CC SMART: SM00352; POU; 1.
CC PROSITE: PS00027; HOMEBOX_1; 1.
CC PROSITE: PS00035; POU_1; 1.
CC PROSITE: PS00465; POU_2; 1.
CC PROSITE: PS00071; HOMEBOX_2; 1.
CC Homeobox; DNA-binding; Transcription regulation; Nuclear protein;
CC Activator; Developmental protein.
CC FT DOMAIN 5 10 POLY-GLN.
CC FT DOMAIN 74 79 POLY-GLU.
CC FT DNA_BIND 290 360 POU.
CC FT DNA_BIND 391 450 HOMEBOX.
CC FT CONFLICT 221 221 V -> GA (IN REF. 2).
CC FT CONFLICT 223 224 RH -> AR (IN REF. 2).
CC FT CONFLICT 248 248 M -> S (IN REF. 4).
CC FT CONFLICT 447 447 K -> N (IN REF. 4).
CC FT CONFLICT 472 474 POA -> RRL (IN REF. 2).
CC FT CONFLICT 475 498 MISSING (IN REF. 2).
CC SEQUENCE 498 AA; 55462 MW; 60F17AF776603974 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 498;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 8 SPES 12
DB 382 SPES 386

```

```

DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sucrose porin precursor.
GN SCRY.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Klebsiella.
OX NCBI_TaxID=573;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1033-5P14 / KAY2026;
RX MEDLINE=91312133; PubMed=1649946;
RT Schmid K., Ebner R., Jahreis K., Lengeler J.W., Tilgemyer F.;
RT "A sugar-specific porin, Scry, is involved in sucrose uptake in
RT enteric bacteria."
RT Mol. Microbiol. 5:941-950(1991).
CC -1- FUNCTION: Porin for sucrose uptake.
CC -1- SUBUNIT: HOMOTRIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
CC -1- DOMAIN: THE C-TERMINUS HELPS TO ANCHOR THE PORIN TO THE OUTER
CC MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE LAMB FAMILY OF PORINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-stb.ch/announce/
CC or send an email to license@isb-stb.ch).
CC -----
CC EMBL: X57401; CAA40657.1; -
CC PIR: S15194; S15194.
CC HSP: P22340; 1A0F.
CC InterPro: IPR003192; Porin_Lamb.
CC Pfam: PF02264; Lamb; 1.
CC Transport; Sugar transport; Outer membrane; Transmembrane; Porin;
CC Signal.
CC CHAIN 1 22 POTENTIAL.
CC FT SIGNAL 23 505 SUCROSE PORIN.
CC FT CHAIN 23 505
CC SEQUENCE 505 AA; 55650 MW; DD38DB241668D5B CRC64;

Query Match 33.3%; Score 5; DB 1; Length 505;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 11 ESRAA 15
DB 51 ESRAA 55

RESULT 107
ID SYR_XYLFA STANDARD; PRT; 506 AA.
AC O9PEB6;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Lysyl-CRNA synthetase (EC 6.1.1.6) (Lysine--tRNA ligase) (LysRS).
GN LYS OR XFI112.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9a5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Brites M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrier H.,
RA Colaudo N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,

```

RL J. Bacteriol. 171:3926-3932(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-1021:
RX MEDLINE-21396507; PubMed-11481430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puhler A., Purnelle B., Ransperger U.,
RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
CC -1- CATALYTIC ACTIVITY: ATP + L-glutamate + tRNA(Glu) = AMP +
CC phosphate + L-glutamyl-tRNA(Glu).
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M27221; AAC35209.1; -.
DR EMBL: AL591792; CAC47482.1; -.
DR PIR: A32886; SYRZET.
DR HSSP: P27000; IGLN.
DR InterPro: IPR004527; Glx_bact.
DR InterPro: IPR000924; Glu_tRNA-synt_1c.
DR InterPro: IPR001412; tRNA-synt_1.
DR Pfam: PF00749; tRNA-synt_1c; 1.
DR PRINTS: PR00987; TRNASYNTHGLU.
DR TIGRPFAMS: TIGR00464; glx_bact; 1.
DR PROSITE: PS00178; AA_tRNA_LIGASE_I; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT SITE 12 22 "HIGH" REGION.
FT SITE 23 257 "RMSKS" REGION.
FT BINDING 256 256 ATP (BY SIMILARITY).
FT CONFLICT 427 427 MISSING (IN REF. 1).
FT CONFLICT 485 485 K -> KO (IN REF. 2).
SQ SEQUENCE 485 AA; 54299 MW; 01897C93708FE978 CRC64;
Query Match 33.3%; Score 5; DB 1; Length 485;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 RSNP 11
Db 51 RSTP 55
RESULT 105
PDM2_DROME
ID PDM2_DROME STANDARD: PRT; 498 AA.
AC P31369; Q24430; Q9VK70;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE POU domain protein 2 (drosophila protein) (Pdm-2) (dPou-28) (dOct2)
DE (molt-mere).
GN PDM2 OR PDM-2 OR DIM OR POU-28 OR OCT2 OR CG1287.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE-92144419; PubMed-1685891;
RA Lloyd A., Sakonju S.;
RT "Characterization of two Drosophila POU domain genes, related to
RT oct-1 and oct-2, and the regulation of their expression patterns.";
RL Mech. Dev. 36:87-102(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-91352045; PubMed-1881906;
RA Dick T., Yang X., Yeo S., Chia W.;
RT "Two closely linked Drosophila POU domain genes are expressed in
RT neuroblasts and sensory elements.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:7645-7649(1991).
RN [3]
RP SEQUENCE OF 68-498 FROM N.A.
RC STRAIN-Oregon-R;
RX MEDLINE-92001544; PubMed-1680380;
RA Billin A.N., Cockrell K.A., Poole S.J.;
RT "Isolation of a family of Drosophila POU domain genes expressed in
RT early development.";
RL Mech. Dev. 34:75-84(1991).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE-Embryo;
RX MEDLINE-92357775; PubMed-1496003;
RA Prakash K., Fang X.D., Engelberg D., Behal A., Parker C.S.;
RT "dOct2, a Drosophila Oct transcription factor that functions in
RT yeast.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:7080-7084(1992).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN-Berkeley;
RX MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Ceolniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Branton R.C., Rogers Y.-H.C., Blazer R.G., Chang M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup I.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gargiellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kenison J.A., Kethum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mitsuhashi N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Schaefer F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RC Science 287:2185-2195(2000).
CC -1- FUNCTION: DNA-BINDING REGULATORY PROTEIN IMPLICATED IN EARLY
CC DEVELOPMENT. INVOLVED IN NEURONAL CELL FATE DECISION. MAY ACT AS
CC AN OCTAMER-DEPENDENT ACTIVATOR OF TRANSCRIPTION. COULD ALSO PLAY

DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Glutamyl-tRNA synthetase (EC 6.1.1.17) (Glutamate--tRNA ligase)
GN (GIURS)
GN GLTX OR H10274.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OC NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McEnney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shiley R., Liu L.-I., Glodok A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudak D.M., Brandon R.C.,
RA Fine L.D., Fitchman J.L., Fuhrmann J.L., Geophagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
Rd.";
RL Science 269:496-512(1995).
CC -1- CATALYTIC ACTIVITY: $\text{ATP} + \text{L-glutamate} + \text{tRNA(Glu)} = \text{AMP} +$
CC $\text{diphosphate} + \text{L-glutamyl-tRNA(Glu)}$.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO CLASS-1 AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U32713; AAC21940.1; -.
DR HSSP: P27000; IGLN.
DR TIGR: H10274; -.
DR InterPro: IPR004527; GLTX_bact.
DR InterPro: IPR000924; GLU_tRNA-synt_1c.
DR InterPro: IPR001412; tRNA-synt_1.
DR Pfam: PF00749; tRNA-synt_1c; 1.
DR PRINTS: PR00987; TRNASYNTHGLU.
DR TIGRFS: TIGR00464; gltx_bact; 1.
DR PROSITE: PS00178; AA_TRNA_LIGASE_1; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT SITE 21 "HIGH" REGION.
FT SITE 248 "KMSK" REGION.
FT BINDING 251 ATP (BY SIMILARITY).
SQ SEQUENCE 480 AA: 54874 MW: 48520396700P CRC64:

Query Match 33.3%; Score 5; DB 1; Length 480;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
DB 60 RSTPE 64

RESULT 103
STE_PASMO
ID STE_PASMO STANDARD: PRT; 480 AA.
AC P57906;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glutamyl-tRNA synthetase (EC 6.1.1.17) (Glutamate--tRNA ligase)

DE (GIURS).
GN GLTX OR PM1115.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OC NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.-J., Zhang Q., Li L.L., Paustian M.L., Whittem T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -1- CATALYTIC ACTIVITY: $\text{ATP} + \text{L-glutamate} + \text{tRNA(Glu)} = \text{AMP} +$
CC $\text{diphosphate} + \text{L-glutamyl-tRNA(Glu)}$.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO CLASS-1 AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE006152; AAK03199.1; -.
DR HSSP: P27000; IGLN.
DR InterPro: IPR004527; GLTX_bact.
DR InterPro: IPR000924; GLU_tRNA-synt_1c.
DR InterPro: IPR001412; tRNA-synt_1.
DR Pfam: PF00749; tRNA-synt_1c; 1.
DR PRINTS: PR00987; TRNASYNTHGLU.
DR TIGRFS: TIGR00464; gltx_bact; 1.
DR PROSITE: PS00178; AA_TRNA_LIGASE_1; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT SITE 21 "HIGH" REGION.
FT SITE 248 "KMSK" REGION.
FT BINDING 251 ATP (BY SIMILARITY).
SQ SEQUENCE 480 AA: 54480 MW: 027A0DA08C4B62DA CRC64:

Query Match 33.3%; Score 5; DB 1; Length 480;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
DB 60 RSTPE 64

RESULT 104
STE_RHIME
ID STE_RHIME STANDARD: PRT; 485 AA.
AC P15189;
DT 01-APR-1990 (Rel. 14, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Glutamyl-tRNA synthetase (EC 6.1.1.17) (Glutamate--tRNA ligase)
GN (GIURS).
GN GLTX OR R02903 OR SMC03172.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OC NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A2;
RX MEDLINE=89291743; PubMed=2661539;
RA Laberge S., Gagnon Y., Bordenau L.M., Lapointe J.;
RT "Cloning and sequencing of the gltx gene, encoding the glutamyl-tRNA
synthetase of Rhizobium meliloti A2.";


```

DR InterPro: IPR004527; Gltx_bact.
DR InterPro: IPR000924; Glu_trna-synt_1c.
DR InterPro: IPR001412; trna-synt_1c.
DR Pfam: PF00749; trna-synt_1c; 1.
DR PRINTS: PR00987; TRNASYNTHGLU.
DR TIGRfams: TIGR00464; gltx_bact; 1.
DR PROSITE: PS00178; AA-trna_LIGASE_I; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
FT SITE 15 25 "HIGH" REGION.
FT SITE 243 247 "KMSKS" REGION.
FT BINDING 246 246 ATP (BY SIMILARITY).
SQ SEQUENCE 471 AA; 52136 MW; 044EBC102AB5FCB9 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPE 11
DB 54 RSTPE 58

RESULT 100
ID SVE_SALTY STANDARD; PRT; 471 AA.
AC 08Z9N7;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Glutamyl-tRNA synthetase (EC 6.1.1.17) (Glutamate--tRNA ligase)
DE (GIURS).
GN GLTX OR STM2415.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SCSJ412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dame M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2."
RL Nature 413:852-856(2001).
CC -1- CATALYTIC ACTIVITY: ATP + L-glutamate + tRNA(Glu) = AMP +
CC diphosphate + L-glutamyl-tRNA(Glu).
CC -1- SUBUNIT: Monomer (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC
CC EMBL: AE008809; AAL21313.1; -
DR StyGene: SG72727; gltx.
DR InterPro: IPR004527; Gltx_bact.
DR InterPro: IPR000924; Glu_trna-synt_1c.
DR InterPro: IPR001412; trna-synt_1c.
DR Pfam: PF00749; trna-synt_1c; 1.
DR PRINTS: PR00987; TRNASYNTHGLU.
DR TIGRfams: TIGR00464; gltx_bact; 1.
DR PROSITE: PS00178; AA-trna_LIGASE_I; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.

```

```

FT SITE 9 19 "HIGH" REGION.
FT SITE 237 241 "KMSKS" REGION.
FT BINDING 240 240 ATP (BY SIMILARITY).
SQ SEQUENCE 471 AA; 53633 MW; F16EB3BBA0CC617 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPE 11
DB 48 RSTPE 52

RESULT 101
ID YH06_YEAST STANDARD; PRT; 471 AA.
AC P38770;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Hypothetical 53.4 kDa protein in SLT2-PUT2 intergenic region.
GN YHR036W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycos.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=94378003; PubMed=8091229;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
RA Du Z., Favell A., Fulton L., Gattung S., Geisel C., Kirsten J.,
RA Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,
RA Latreille P., Louis E.J., Macri C., Mardis E., Meneses S., Mouser L.,
RA Nhan M., Rifkin L., Riles L., St. Peter H., Treviski E., Vaughan K.,
RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
RA Vaudin M.;
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
RT VIII."
RL Science 265:2077-2082(1994).
CC -1- SIMILARITY: SOME, TO YEAST YGL247W.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC
CC EMBL: 000062; AAB68908.1; -
DR PIR: S46739; S46739.
DR SGD: S0001078; YHR036W.
DR Hypothetical protein; Transmembrane.
KW TRANSMEM 300 320 POTENTIAL.
FT TRANSMEM 408 428 POTENTIAL.
SQ SEQUENCE 471 AA; 53385 MW; 732209C74DE3463 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PHRST 9
DB 140 PHRST 144

RESULT 102
ID SVE_HAELN STANDARD; PRT; 480 AA.
AC P43818;
DT 01-NOV-1995 (Rel. 32, Created)

```



```

Db      48 RSTPE 52

|||||
RESULT 98
SYE_ECOLI STANDARD: PRT: 471 AA.
AC P04805;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glutamyl-tRNA synthetase (EC 6.1.1.17) (Glutamate--tRNA ligase)
GN (GltR).
OS Escherichia coli.
OC Escherichia coli.
CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
CX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=66278132; PubMed=3015933;
RA Breton R., Santacon H., Papayannopoulos I., Blemann K., Lapointe J.;
RT "Glutamyl-tRNA synthetase of Escherichia coli. Isolation and primary
RT structure of the gltx gene and homology with other aminoacyl-tRNA
RT synthetases."
RL J. Biol. Chem. 261:10610-10617(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Valdes J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=K12;
RX MEDLINE=97349980; PubMed=9205837;
RA Yamamoto Y., Alba H., Baba T., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsunashi N.,
RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nishimoto H.,
RA Oshima T., Oyama S., Saito N., Sampei G., Satoh Y., Sivasubram S.,
RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA Yamagata S., Horituchi T.;
RT "Construction of a contiguous 874-kb sequence of the Escherichia coli
RT K-12 genome corresponding to 50.0-68.8 min on the linkage map and
RT analysis of its sequence features."
RL DNA Res. 4:91-113(1997).
RN [4]
RP SEQUENCE OF 1-57 FROM N.A.
RX MEDLINE=90355200; PubMed=2201777;
RA Brun V., Santacon H., Breton R., Lapointe J.;
RT "Closely spaced and divergent promoters for an aminoacyl-tRNA
RT synthetase gene and a tRNA operon in Escherichia coli.
RT Transcriptional and post-transcriptional regulation of gltx, valU
RT and alaW."
RL J. Mol. Biol. 214:845-864(1990).
CC -1- CATALYTIC ACTIVITY: ATP + L-glutamate + tRNA(Glu) = AMP +
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- MISCELLANEOUS: THIS IS THE SMALLEST AMINOACYL-TRNA SYNTHETASE OF
CC E. COLI. IT DOES NOT BIND GLUTAMATE IN THE ABSENCE OF COGNATE tRNA,
CC WHICH IS THEREFORE REQUIRED FOR ACTIVATION OF THE AMINO ACID
CC SUBSTRATE.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; X63976; CAA45391.1; -
CC DR EMBL; M13687; AAA65715.1; -
CC DR EMBL; AE000328; AAC75457.1; -
CC DR EMBL; D90869; BAA16272.1; -
CC DR EMBL; X55737; CAA39269.1; -
CC DR PIR; A25956; SYRDET.
CC HSSP; P27000; IGIN.
CC DR ECODBASE; F047.8; 6TH EDITION.
CC DR EcGene; EG10407; gltx.
CC DR InterPro: IPR004527; GltX_bact.
CC DR InterPro: IPR000924; Glt_tRNA-synt_1c.
CC DR InterPro: IPR001412; tRNA-synt_1.
CC DR Pfam; PF00749; tRNA-synt_1c; 1.
CC DR PRINTS; PR00987; TRNASYNTGLD.
CC DR TIGRFAMS; TIGR00464; gltx_bact; 1.
CC DR PROSITE; PS00178; AA-TRNA_LIGASE_I; 1.
CC KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
CC Complete proteome.
CC FT SITE 9 "HIGH" REGION.
CC FT SITE 237 "KMSKS" REGION.
CC FT BINDING 240 ATP (BY SIMILARITY).
CC FT BINDING 240 ATP (BY SIMILARITY).
CC SO SEQUENCE 471 AA; 53815 MW; 8264A799E5383398 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 471;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 7 RSTPE 11
Db 48 RSTPE 52

RESULT 99
SYE_RHOSH STANDARD: PRT: 471 AA.
AC Q9ZFA3;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glutamyl-tRNA synthetase (EC 6.1.1.17) (Glutamate--tRNA ligase)
GN (GltR).
OS Rhodospirillum rubrum.
OC Bacteria; Proteobacteria; alpha subdivision; Rhodospirillum rubrum;
CC Rhodospirillum rubrum.
CX NCBI_TaxID=1063;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 17023 / 2.4.1 / NCIB 8253 / DSM 158;
RA Simmons A.E., Mackenzie R.C., Kaplan S.;
RT "Tri-ling up paradigms: distribution of tryptophan biosynthesis genes
RT in Rhodospirillum rubrum 2.4.1."
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + L-glutamate + tRNA(Glu) = AMP +
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; AF108766; AAD09123.1; -
CC HSSP; P27000; IGIN.

```

DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE POU domain protein 2 (didymous protein) (Pdm-2) (dPOU-28) (dOCT2)
 DE (mitf-mer) (Fragment).
 GN PDM2 OR PDM-2 OR DIM OR POU-28 OR OCT2.
 OS Drosophila virilis (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachyera;
 OC Muscomorpha; Bphytroidea; Drosophilidae; Drosophila.
 NC NCBI_TaxID=7244;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP TISSUE=Embryo;
 RX MEDLINE=95267688; PubMed=7748782;
 RA Poole S.J.;
 RT "Conservation of complex expression domains of the pdm-2 POU domain
 gene between Drosophila virilis and Drosophila melanogaster.";
 RT Mech. Dev. 49:107-116(1995).
 CC -1- FUNCTION: DNA-BINDING REGULATORY PROTEIN IMPLICATED IN EARLY
 DEVELOPMENT. INVOLVED IN NEURONAL CELL FATE DECISION. MAY ACT AS
 AN OCTAMER-DEPENDENT ACTIVATOR OF TRANSCRIPTION (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED PRIMARILY DURING THE FIRST HALF OF
 EMBRYOGENESIS. INITIAL EXPRESSION IN CELLULAR BLASTODERM STAGE,
 THEN IN ECTODERMAL STRIPES DURING GERMBAND EXTENSION. BROAD
 EXPRESSION IN THE NEUROECTODERM FOLLOWED BY LIMITATION TO DISCRETE
 SUBSETS OF CNS CELLS. AND EXPRESSION IN SPECIFIC PNS NEURONS AND
 SUPPORT CELLS.
 CC -1- SIMILARITY: BELONGS TO THE POU TRANSCRIPTION FACTOR FAMILY.
 CC -1- CLASS-2 SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 1 HOMEOBOX DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation-
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U14723; AAA21584.1; -
 DR HSP: P14859; 1OCT.
 DR Flybase; FBgn0020449; Dvir/pdm2.
 DR InterPro; IPR001356; Homeobox.
 DR InterPro; IPR000327; POU_domain.
 DR Pfam; PF00046; homeobox; 1.
 DR Pfam; PF00157; pou; 1.
 DR ProDom; PD000010; Homeobox; 1.
 DR ProDom; PD000583; POU_domain; 1.
 DR SMART; SM00389; HOX; 1.
 DR SMART; SM00352; POU; 1.
 DR PROSITE; PS00027; HOMEOBOX_1; 1.
 DR PROSITE; PS50071; HOMEOBOX_2; 1.
 DR PROSITE; PS00035; POU_1; 1.
 DR PROSITE; PS00465; POU_2; 1.
 DR Homeobox; DNA-binding; Transcription regulation; Nuclear protein;
 KW Activator; Developmental protein.
 FT NON_TER 1
 FT DOMAIN 1
 FT DOMAIN 6 12 POLY-GLU.
 FT DOMAIN 70 73 POLY-ALA.
 FT DOMAIN 218 223 POLY-GLN.
 FT DOMAIN 361 331 POU.
 FT DNA_BIND 362 421 HOMEOBOX.
 SO SEQUENCE 471 AA; 51225 MW; BDD27F895EC2428C CRC64;

RESULT 97
 SVE_ECO57
 ID SVE_ECO57 STANDARD; PRT; 471 AA.
 AC 08XBN2;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Glutamyl-tRNA synthetase (EC 6.1.1.17) (Glutamate--tRNA ligase)
 DE (GLURS).
 GN GLTX OR Z3665 OR ECS3278.
 OS Escherichia coli O157:H7.
 OS Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 NC NCBI_TaxID=83334;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G., III, Burland V., Mau B., Glasner J.D.,
 Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 Grobeck E.J., Davis N.W., Llm A., Dimalanta E.T., Potamousis K.,
 Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
 RT Nature 409:529-533(2001).
 RN [2]
 RC SEQUENCE FROM N.A.
 RP STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 Han C.-G., Ohsudo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 Ikeda T., Takami H., Honda T., Sasekawa C., Ogasawara N., Yasunaga T.,
 Kuida S., Shiba T., Hattori M., Shitagawa H.;
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 O157:H7 and genomic comparison with a laboratory strain K-12.";
 RT DNA Res. 8:11-22(2001).
 CC -1- CATALYTIC ACTIVITY: ATP + L-glutamate + tRNA(Glu) -> AMP +
 diphosphate + L-glutamyl-tRNA(Glu).
 CC -1- SUBUNIT: Monomer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO CLASS-1 AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation-
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE005471; AAG57524.1; -
 DR EMBL: AP002561; BAB36701.1; -
 DR InterPro; IPR004527; GLTX_dact.
 DR InterPro; IPR000924; Glu_tRNA-synt_1c.
 DR InterPro; IPR001412; tRNA-synt_1.
 DR Pfam; PF00749; tRNA-synt_1c; 1.
 DR PRINTS; PR00967; TRNASINTHGLU.
 DR TIGRFAMS; TIGR00464; gltx_dact; 1.
 DR PROSITE; PS00178; AA_TRNA_LIGASE_1; 1.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 CC Complete proteome.
 FT SITE 9 19 "HIGH" REGION.
 FT BINDING 237 241 "KMSKS" REGION.
 FT SITE 240 240 ATP (BY SIMILARITY).
 SO SEQUENCE 471 AA; 53755 MW; FCSA227F3CEFF2592.CRC64;

Query Match 33.3%; Score 5; DB 1; Length 471;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 8 STPES 12
 Db 353 STPES 357
 7 RSTPE 11

FT CONFLICT 159 159 I -> L (IN REF. 3).
 SQ SEQUENCE 465 AA; 52088 MW; 8C85ED079A52D61 CRC64;
 Query Match 33.3%; Score 5; DB 1; Length 465;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
 |||||
 DB 63 RSTPE 67

RESULT 95
 HKX4_RAT
 ID HKX4_RAT STANDARD: PRT; 465 AA.
 AC P17712; P17711;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hexokinase D (Ec 2.7.1.1) (Hexokinase type IV) (HK IV) (HK4)
 DE (Glucokinase).
 GN GCK.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_Taxid=10116;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 3), AND PARTIAL SEQUENCE.
 RC TISSUE=Liver;
 RX MEDLINE=89079677; PubMed=2909525;
 RA Andeone T.L., Printz R.L., Plikis S.J., Magnuson M.A., Granner D.K.;
 RT "The amino acid sequence of rat liver glucokinase deduced from cloned
 RT cDNA.";
 RL J. Biol. Chem. 264:363-369(1989).
 RN [2]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RC TISSUE=Liver;
 RX MEDLINE=90372925; PubMed=2396986;
 RA Hayzer D.V., Iynedjian P.B.;
 RT "Alternative splicing of glucokinase mRNA in rat liver.";
 RL Biochem. J. 270:261-263(1990).
 RN [3]
 RP SEQUENCE OF 1-15 FROM N.A. (ISOFORM 1).
 RC TISSUE=Pancreas;
 RX MEDLINE=89380186; PubMed=2550428;
 RA Magnuson M.A., Shelton K.D.;
 RT "An alternate promoter in the glucokinase gene is active in the
 RT pancreatic beta cell.";
 RL J. Biol. Chem. 264:15936-15942(1989).
 RN [4]
 RP SEQUENCE OF 1-15 FROM N.A. (ISOFORM 3).
 RX MEDLINE=89296899; PubMed=2662183;
 RA Magnuson M.A., Andeone T.L., Printz R.L., Koch S., Granner D.K.;
 RT "Rat glucokinase gene: structure and regulation by insulin.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:4838-4842(1989).
 RN [5]
 RP SEQUENCE OF 1-15 FROM N.A. (ISOFORM 3).
 RX MEDLINE=90073655; PubMed=2590200;
 RA Noguchi T., Takenaka M., Yamada K., Matsuda T., Hashimoto M.,
 RA Tanaka T.;
 RT "Characterization of the 5' flanking region of rat glucokinase gene.";
 RL Biochem. Biophys. Res. Commun. 164:1247-1252(1989).
 RN [6]
 RP SEQUENCE OF 1-166 FROM N.A. (ISOFORM 1).
 RX MEDLINE=91154262; PubMed=1999433;
 RA Hughes S.D., Quade C., Milburn J.L., Cassidy L., Newgard C.B.;
 RT "Expression of normal and novel glucokinase mRNAs in anterior
 RT pituitary and islet cells.";
 RL J. Biol. Chem. 266:4521-4530(1991).
 CC -1- FUNCTION: CATALYZES THE INITIAL STEP IN UTILIZATION OF GLUCOSE BY
 CC THE BETA-CELL AND LIVER AT PHYSIOLOGICAL GLUCOSE CONCENTRATION.
 CC PANCREATIC GLUCOKINASE PLAYS AN IMPORTANT ROLE IN MODULATING
 CC INSULIN SECRETION. HEPATIC GLUCOKINASE HELPS TO FACILITATE THE

CC UPTAKE AND CONVERSION OF GLUCOSE BY ACTING AS AN INSULIN-SENSITIVE
 CC DETERMINANT OF HEPATIC GLUCOSE USAGE.
 CC -1- CATALYTIC ACTIVITY: ATP + D-hexose = ADP + D-hexose 6-phosphate.
 CC -1- ENZYME REGULATION: THE USE OF ALTERNATIVE PROMOTERS APPARENTLY
 CC ENABLES THE TYPE IV HEXOKINASE GENE TO BE REGULATED BY INSULIN
 CC IN THE LIVER AND GLUCOSE IN THE BETA CELL. THIS MAY CONSTITUTE AN
 CC IMPORTANT FEEDBACK LOOP FOR MAINTAINING GLUCOSE HOMEOSTASIS.
 CC -1- ALTERNATIVE PRODUCTS: 3 isoforms; 1 (shown here), 2 and 3; are
 CC produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: PANCREAS (ISOFORMS 1 AND 2) AND LIVER (ISOFORM
 CC 3).
 CC -1- MISCELLANEOUS: IN VERTEBRATES THERE ARE FOUR MAJOR GLUCOSE-
 CC PHOSPHORYLATING ISOENZYMES, DESIGNATED HEXOKINASE I, II, III AND
 CC IV (GLUCOKINASE).
 CC -1- SIMILARITY: BELONGS TO THE HEXOKINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch)
 CC -----

DR EMBL; J04218; AAA41229.1; -;
 DR EMBL; M24952; AAA41230.1; -;
 DR EMBL; M24943; AAA41230.1; JOINED.
 DR EMBL; M24944; AAA41230.1; JOINED.
 DR EMBL; M24945; AAA41230.1; JOINED.
 DR EMBL; M24946; AAA41230.1; JOINED.
 DR EMBL; M24948; AAA41230.1; JOINED.
 DR EMBL; M24949; AAA41230.1; JOINED.
 DR EMBL; M24950; AAA41230.1; JOINED.
 DR EMBL; M24951; AAA41230.1; JOINED.
 DR EMBL; M25806; AAA41238.1; -;
 DR EMBL; M25807; AAA41239.1; -;
 DR EMBL; M58789; AAA41236.1; -;
 DR EMBL; X53588; CAA37657.1; -;
 DR EMBL; X53590; CAA37660.1; -;
 DR EMBL; M30770; AAA41231.1; -;
 DR PIR; A31810; A31810.
 DR PIR; S11465; S11465.
 DR PIR; S12062; S12062.
 DR HSSP; P35557; IGUK.
 DR InterPro; IPR001312; Hexokinase.
 DR Pfam; PF00349; hexokinase; 1.
 DR Pfam; PF03727; hexokinase2; 1.
 DR PRINTS; PR00475; HEXOKINASE.
 DR PRODOM; PD001109; Hexokinase; 1.
 DR PROSITE; PS00378; HEXOKINASES; 1.
 DR Transferrase; Kinase; Glycolysis; ATP-binding; Alternative splicing.
 KW NP_BIND 78
 FT BINDING 104 104 ATP (POTENTIAL).
 FT DOMAIN 145 171 GLUCOSE-BINDING (POTENTIAL).
 FT VARSPLIC 1 15 MLDNRARREATKKEK -> MAMDTRCGAQLITL (IN
 FT VARSPLIC 1 15 ISOFORM 3).
 FT FT
 FT VARSPLIC 122 138 MISSING (IN ISOFORM 2).
 SQ SEQUENCE 465 AA; 52116 MW; 65B36673DE9297BE CRC64;

Query Match 33.3%; Score 5; DB 1; Length 465;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
 |||||
 DB 63 RSTPE 67

RESULT 96
 PMW2_DROVI
 ID PMW2_DROVI STANDARD: PRT; 471 AA.
 AC Q24705;
 DT 15-DEC-1998 (Rel. 37, Created)

RL Hum. Mutat. 12:136-136(1998).
 RP [18]
 RP VARIANTS MODY2 H-108; S-150; T-259; R-299; Y-382; T-384 AND C-392.
 RX MEDLINE=98324778; PubMed=9662401.
 RA Hattersley A.T., Beards F., Ballantyne E., Appleton M., Harvey R.,
 RA Ellard S.,
 RT "Mutations in the glucokinase gene of the fetus result in reduced
 RT birth weight."
 RL Nat. Genet. 19:268-270(1998).
 RN [19]
 RP VARIANT HYPERINSULINISM MET-455.
 RX MEDLINE=98084322; PubMed=9435328.
 RA Glaser B., Kesavan P., Heyman M., Davis E., Cuesta A., Buchs A.,
 RA Stanley C.A., Thornton P.S., Permutt M.A., Matschinsky F.M.,
 RA Herold K.C.;
 RT "Familial hyperinsulinism caused by an activating glucokinase
 RT mutation."
 RL New Engl. J. Med. 338:226-230(1998).
 RN [20]
 RP VARIANTS MODY2 THR-110; ASP-119 AND VAL-385.
 RX NG M.C.Y., Cockburn B.N., Lindner T.H., Yeung V.T.F., Chow C.-C.,
 RA So W.-Y., Li J.K.Y., Lo Y.M.D., Lee Z.S.K., Cockram C.S.,
 RA Critchley J.A.J.H., Bell G.I., Chan J.C.N.;
 RT "Molecular genetics of diabetes mellitus in Chinese subjects:
 RT identification of mutations in glucokinase and hepatocyte nuclear
 RT factor-1alpha genes in patients with early-onset type 2 diabetes
 RT mellitus/MODY."
 RL Diabet. Med. 16:956-963(1999).
 RN [21]
 RP VARIANT MODY2 PRO-164.
 RX MEDLINE=20560768; PubMed=11106831;
 RA Nam J.H., Lee H.C., Kim Y.H., Cha B.S., Song Y.D., Lim S.K., Kim K.R.,
 RA Huh K.B.;
 RT "Identification of glucokinase mutation in subjects with post-renal
 RT transplantation diabetes mellitus."
 RL Diabetes Res. Clin. Pract. 50:169-176(2000).
 RN [22]
 RP VARIANTS MODY2 LYS-210 AND MET-228.
 RX MEDLINE=21245203; PubMed=11372010;
 RA Njolstad P.R., Soevik O., Cuesta-Munoz A., Bjorkhaug L., Massa O.,
 RA Barbelli F., Undlien D.E., Shiota C., Magnuson M.A., Molven A.,
 RA Matschinsky F.M., Bell G.I.;
 RT "Neonatal diabetes mellitus due to complete glucokinase deficiency."
 RL New Engl. J. Med. 344:1588-1592(2001).
 CC -1- FUNCTION: CATALYZES THE INITIAL STEP IN UTILIZATION OF GLUCOSE BY
 CC THE BETA-CELL AND LIVER AT PHYSIOLOGICAL GLUCOSE CONCENTRATION.
 CC GLUCOKINASE HAS A HIGH KM FOR GLUCOSE, AND SO IT IS EFFECTIVE ONLY
 CC WHEN GLUCOSE IS ABUNDANT. THE ROLE OF GCK AND ISO IT IS EFFECTIVE ONLY
 CC THE SYNTHESIS OF GLYCOGEN. PANCREATIC GLUCOKINASE PLAYS AN
 CC IMPORTANT ROLE IN MODULATING INSULIN SECRETION.

Query Match 33.3%; Score 5; DB 1; Length 465;
 Best Local Similarity 100.0%; Pred. No. 1,8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 7 RSTPE 11
 Db 63 RSTPE 67

RESULT 94
 HXK4 MOUSE
 ID HXK4 MOUSE STANDARD; PRT; 465 AA.
 AC P52792; P52791;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hexokinase D (EC 2.7.1.1) (Hexokinase type IV) (HK IV) (HK4)
 DE (Glucokinase).
 GN GCK OR GK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=96121388; PubMed=8575769.
 RA Ishimura-Oka K., Nakamura M., Chu M.J., Sullivan M., Chan L.,
 RA Oka K.;
 RT "Partial structure of the mouse glucokinase gene."
 RL Genomics 29:751-754(1995).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC STRAIN=129/SV; TISSUE=Liver;
 RX MEDLINE=96121387; PubMed=8575768;
 RA Postle C., Niswender K.D., Decaux J.F., Shelton K.D., Goubot B.,
 RA Petherph C.C., Graner D.K., Girard J., Magnuson M.A.;
 RT "Cloning and characterization of the mouse glucokinase gene locus and
 RT identification of distal liver-specific Dnase I hypersensitive
 RT sites."
 RL Genomics 29:740-750(1995).
 RN [3]
 RP SEQUENCE OF 1-166 FROM N.A. (ISOFORM 1).
 RC TISSUE=Pancreas;
 RX MEDLINE=91154262; PubMed=1999433;
 RA Hughes S.D., Quade C., Milburn J.L., Cassidy L., Newgard C.B.;
 RT "Expression of normal and novel glucokinase mRNAs in anterior
 RT pituitary and islet cells."
 RL J. Biol. Chem. 266:4521-4530(1991).
 CC -1- FUNCTION: CATALYZES THE INITIAL STEP IN UTILIZATION OF GLUCOSE BY
 CC THE BETA-CELL AND LIVER AT PHYSIOLOGICAL GLUCOSE CONCENTRATION.
 CC PANCREATIC GLUCOKINASE PLAYS AN IMPORTANT ROLE IN MODULATING
 CC INSULIN SECRETION. HEPATIC GLUCOKINASE HELPS TO FACILITATE THE
 CC UPTAKE AND CONVERSION OF GLUCOSE BY ACTING AS AN INSULIN-SENSITIVE
 CC DETERMINANT OF HEPATIC GLUCOSE USAGE.
 CC -1- CATALYTIC ACTIVITY: ATP + D-hexose -> ADP + D-hexose 6-phosphate.
 CC -1- ENZYME REGULATION: THE USE OF ALTERNATIVE PROMOTERS APPARENTLY
 CC ENABLES THE TYPE IV HEXOKINASE GENE TO BE REGULATED BY INSULIN
 CC IN THE LIVER AND GLUCOSE IN THE BETA CELL. THIS MAY CONSTITUTE AN
 CC IMPORTANT FEEDBACK LOOP FOR MAINTAINING GLUCOSE HOMEOSTASIS.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
 CC produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: PANCREAS, ANTERIOR PITUITARY (ISOFORM 1) AND
 CC LIVER (ISOFORM 2).
 CC -1- MISCELLANEOUS: IN VERTEBRATES THERE ARE FOUR MAJOR GLUCOSE-
 CC PHOSPHORYLATING ISOENZYMES, DESIGNATED HEXOKINASE I, II, III AND
 CC IV (GLUCOKINASE).
 CC -1- SIMILARITY: BELONGS TO THE HEXOKINASE FAMILY.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL: L38990; AAB00360.1; -
 EMBL: I41631; AAC42074.1; -
 EMBL: M58755; AAA37703.1; -
 DR HSSP: P35557; IGLK.
 DR MGI: 95724; GK.
 DR InterPro: IPR001312; Hexokinase.
 DR Pfam: PF00349; hexokinase.1.
 DR Pfam: PF03727; hexokinase2.1.
 DR PRINTS: PR00475; HEXOKINASE.
 DR PRODOM: PD001109; Hexokinase; 1.
 DR PROSITE: PS00378; HEXOKINASES; 1.
 DR TRANSFERASE: Kinase; Glycolysis;
 FT NP_BIND 78 83 ATP (POTENTIAL).
 FT BINDING 104 104 ATP (POTENTIAL).
 FT DOMAIN 145 171 GLUCOSE-BINDING (POTENTIAL).
 FT VARSPLIC 1 15 MUDDRAMENTYKKK -> MAVDITRGASQSLTL (IN
 FT ISOFORM 2).
 FT F -> L (IN REF. 3).
 FT CONFLICT 133 133

DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hexokinase D (EC 2.7.1.1) (Hexokinase type IV) (HK IV) (HK4)
 DE (Glucokinase).
 GN GCK.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-92375100; PubMed-1354840;
 RA Tanizawa Y., Matsutani A., Chiu K.C., Permutt M.A.;
 RT "Human glucokinase gene: isolation, structural characterization, and
 RT identification of a microsatellite repeat polymorphism.";
 RL Mol. Endocrinol. 6:1070-1081(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-91334452; PubMed-1871135;
 RA Tanizawa Y., Koranyi L.I., Weiling C.M., Permutt M.A.;
 RT "Human liver glucokinase gene: cloning and sequence determination of
 RT two alternatively spliced cDNAs.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:7294-7297(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX TISSUE-Pancreas;
 MEDLINE-92380355; PubMed-1511800;
 RA Nishi S., Stoffel M., Xiang K.S., Shows T.B., Bell G.I., Takeda J.;
 RT "Human pancreatic beta-cell glucokinase: cDNA sequence and
 RT localization of the polymorphic gene to chromosome 7, band p13.";
 RL Diabetologia 35:743-747(1992).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX TISSUE-Pancreas;
 MEDLINE-92307138; PubMed-1612194;
 RA Koranyi L.I., Tanizawa Y., Weiling C.M., Rabin D.U., Permutt M.A.;
 RT "Human islet glucokinase gene: isolation and sequence analysis of
 RT full-length cDNA.";
 RL Diabetes 41:807-811(1992).
 RN [5]
 RP SEQUENCE FROM N.A., AND VARIANTS MODY2 MET-228 AND ARG-261.
 RX TISSUE-Pancreas;
 MEDLINE-92366529; PubMed-1502186;
 RA Stoffel M., Froguel P., Takeda J., Zouali H., Vionnet N., Nishi S.,
 RA Weber I.T., Harrison R.W., Pilakis S.J., Lesage S., Vaxillaire M.,
 RA Velho G., Sun F., Irls F., Passa P., Cohen D., Bell G.I.;
 RT "Human glucokinase gene: isolation, characterization, and
 RT identification of two missense mutations linked to early-onset
 RT non-insulin-dependent (type 2) diabetes mellitus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:7698-7702(1992).
 RN [6]
 RP ERRATUM.
 RA Stoffel M., Froguel P., Takeda J., Zouali H., Vionnet N., Nishi S.,
 RA Weber I.T., Harrison R.W., Pilakis S.J., Lesage S., Vaxillaire M.,
 RA Velho G., Sun F., Irls F., Passa P., Cohen D., Bell G.I.;
 RL Proc. Natl. Acad. Sci. U.S.A. 89:10562-10562(1992).
 RN [7]
 RP SEQUENCE FROM N.A., VARIANT THR-107, AND VARIANT MODY2 ARG-261.
 RX MEDLINE-93100400; PubMed-1464666;
 RA Sakura H., Eto K., Kadowaki H., Shimokawa K., Ueno H., Koda N.,
 RA Fukushima Y., Akanuma Y., Yazaki Y., Kadowaki T.,
 RT "Structure of the human glucokinase gene and identification of a
 RT missense mutation in a Japanese patient with early-onset non-insulin-
 RT dependent diabetes mellitus.";
 RL J. Clin. Endocrinol. Metab. 75:1571-1573(1992).
 RN [8]
 RP SEQUENCE FROM N.A.
 RX TISSUE-Lung;
 RA Strusberg R.;
 RN Submitted (Jan-2001) to the EMBL/GenBank/DBJ databases.
 RN [9]
 RP 3D-STRUCTURE MODELING.
 RX MEDLINE-94252471; PubMed-8194664;

RA St Charles R., Harrison R.W., Bell G.I., Pilakis S.J., Weber I.T.;
 RT "Molecular model of human beta-cell glucokinase built by analogy to
 RT the crystal structure of yeast hexokinase B.";
 RL Diabetes 43:784-791(1994).
 RN [10]
 RP VARIANT MODY2 ARG-299.
 RX MEDLINE-93265142; PubMed-1303265;
 RA Stoffel M., Patel P., Lo Y.-M.D., Hattersley A.T., Lucassen A.M.,
 RA Page R., Bell J.I., Bell G.I., Turner R.C., Wainscoat J.S.;
 RT "Missense glucokinase mutation in maturity-onset diabetes of the
 RT young and mutation screening in late-onset diabetes.";
 RL Nat. Genet. 2:153-156(1992).
 RN [11]
 RP VARIANT THR-11.
 RX MEDLINE-93202338; PubMed-8454109;
 RA Chiu K.C., Tanizawa Y., Permutt M.A.;
 RT "Glucokinase gene variants in the common form of NIDDM.";
 RL Diabetes 42:579-582(1993).
 RN [12]
 RP VARIANT MODY2 PRO-131.
 RX PubMed-8495817;
 RA Stoffel M., Bell K.L., Blackburn C.L., Powell K.L., Seo T.S.,
 RA Takeda J., Vionnet N., Xiang K.-S., Gidh-Jain M., Pilakis S.J.,
 RA Ober C., Bell G.I.;
 RT "Identification of glucokinase mutations in subjects with gestational
 RT diabetes mellitus.";
 RL Diabetes 42:937-940(1993).
 RN [13]
 RP VARIANT N-4, AND VARIANTS MODY2 K-70; P-131; T-188; R-257 AND E-414.
 RX MEDLINE-93315503; PubMed-8325892;
 RA Takeda J., Gidh-Jain M., Xu L.Z., Froguel P., Velho G., Vaxillaire M.,
 RA Cohen D., Shimada F., Makino H., Nishi S., Stoffel M., Vionnet N.,
 RA St Charles R., Harrison R.W., Weber I.T., Bell G.I., Pilakis S.J.;
 RT "Structure/function studies of human beta-cell glucokinase. Enzymatic
 RT properties of a sequence polymorphism, mutations associated with
 RT diabetes, and other site-directed mutants.";
 RL J. Biol. Chem. 268:15200-15204(1993).
 RN [14]
 RP CHARACTERIZATION OF MODY2 VARIANTS.
 RX MEDLINE-93189611; PubMed-8446612;
 RA Gidh-Jain M., Takeda J., Xu L.Z., Lange A.J., Vionnet N., Stoffel M.,
 RA Froguel P., Velho G., Sun D., Cohen D., Patel P., Lo Y.-M.D.,
 RA Hattersley A.T., Lutchman H., Wedell A., St Charles R., Harrison R.W.,
 RA Weber I.T., Bell G.I., Pilakis S.J.;
 RT "Glucokinase mutations associated with non-insulin-dependent (type 2)
 RT diabetes mellitus have decreased enzymatic activity: implications for
 RT structure/function relationships.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:1932-1936(1993).
 RN [15]
 RP VARIANTS MODY2 TRP-36; MET-209 AND GLU-261.
 RX MEDLINE-94222222; PubMed-8168652;
 RA Hager J., Blanche H., Sun F., Vionnet N., Vaxillaire M., Poller M.,
 RA Cohen D., Czerlichow P., Velho G., Robert J.-J., Cohen N., Froguel P.;
 RT "Six mutations in the glucokinase gene identified in MODY by using a
 RT nonradioactive sensitive screening technique.";
 RL Diabetes 43:730-733(1994).
 RN [16]
 RP VARIANTS MODY2.
 RX MEDLINE-97201951; PubMed-9049484;
 RA Velho G., Blanche H., Vaxillaire M., Bellanne-Chanelot C.,
 RA Parini V.C., Timsit J., Passa P., Deschamps I., Robert J.-J.,
 RA Weber I.T., Marotta D., Pilakis S.J., Lipkind G.M., Bell G.I.;
 RT "Identification of 14 new glucokinase mutations and description of the
 RT clinical profile of 42 MODY-2 families.";
 RL Diabetologia 40:217-224(1997).
 RN [17]
 RP VARIANTS MODY2 SER-80; LYS-221 AND CYS-227.
 RX Guazzini B., Gaffi D., Mainieri D., Multari G., Cordera R.,
 RA Bertolini S., Pozza G., Meschi F., Barbetti F.;
 RT "Three novel missense mutations in the glucokinase gene (G80S; E221K;
 RT G227C) in Italian subjects with maturity-onset diabetes of the young
 RT (MODY).";

PROSITE: PS00035; POU.1; 1.
DR PROSITE: PS00465; POU.2; 1.
KM Transcription regulation; Homeobox; DNA-binding; Nuclear protein;
KW Alternative splicing; Developmental protein.
FT DOMAIN 84 98 POLY-ALA.
FT DOMAIN 165 169 POLY-GLN.
FT DOMAIN 211 222 POLY-HIS.
FT DOMAIN 241 312 POU.
FT DNA_BIND 330 389 HOMEBOX.
FT VARSPIC 407 443 GHFLVDYLIKASITGSEPGDQRYTTSSFTQYLAH ->
SQ SEQUENCE 443 AA; 47453 MW; 5F2A6DC6B833F5C CRC64;
VSADTPPSMDCKRMSET (IN SHORT ISOFORM).
Query Match 33.3%; Score 5; DB 1; Length 443;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HLSPH 6
DB 139 HLGPH 143
RESULT 91
VNS1_BMDNV STANDARD; PRT; 455 AA.
ID VNS1_BMDNV
AC P05840;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Putative nonstructural protein (ORF1).
OS Bombyx densonucleosis virus (BMDNV) (Silkworm densovirus).
OC Viruses; ssDNA viruses; Parvoviridae; Densovirus; Iteavirus.
OX NCBI_TaxID=10809;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Isolate INA;
RX MEDLINE=91021065; Pubmed=2219739;
RA Bando H., Choi H., Ito Y., Kawase S.;
RT "terminal structure of a Densovirus implies a hairpin transfer
RT replication which is similar to the model for AAV.";
RL Virology 179:57-63(1990).
[2]
RN SEQUENCE OF 26-455 FROM N.A.
RP STRAIN=Isolate INA;
RX MEDLINE=87112952; Pubmed=3027382;
RA Bando H., Kusuda J., Gojobori T., Maruyama T., Kawase S.;
RT "organization and nucleotide sequence of a densovirus genome imply a
RT host-dependent evolution of the parvoviruses.";
RL J. Virol. 61:553-560(1987).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M60583; AAA42973.1; -
DR EMBL: M15123; AAA67696.1; -
DR PIR: A26796; UYVPL.
KW Nonstructural protein.
SQ SEQUENCE 455 AA; 52488 MW; 0E5466B06C775A04 CRC64;
Query Match 33.3%; Score 5; DB 1; Length 455;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 RSTPE 11
DB 382 RSTPE 386

RESULT 92
GBAL_DROME STANDARD; PRT; 457 AA.
ID GBAL_DROME
AC P25157;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Guanine nucleotide-binding protein, alpha subunit homolog (Protein
DE concertina).
DE CTA OR CTR.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Oregon-R;
RX MEDLINE=91105860; Pubmed=1839050;
RA Parks S., Wieschaus E.;
RT "The Drosophila gastrulation gene concertina encodes a G alpha-like
RT protein.";
RL Cell 64:447-458(1991).
CC -1- FUNCTION: MAY PLAY A ROLE IN A SIGNAL TRANSDUCTION PATHWAY USED
CC DURING GASTRULATION. REQUIRED SPECIFICALLY FOR THE VENTRAL FURROW
CC AND POSTERIOR MIDGUT INVAGINATIONS, WHERE IT IS NECESSARY FOR
CC COORDINATING CELL SHAPE CHANGES.
CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC SIGNALING SYSTEMS.
CC -1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC -1- TISSUE SPECIFICITY: DISTRIBUTED UNIFORMLY.
CC -1- DEVELOPMENTAL STAGE: GASTRULATION.
CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY, SUBFAMILY 4 (G(12)).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M64285; AAA82939.1; -
DR PIR: A38567; A38567.
DR HSSP: P04896; IAZT.
DR FLYBASE: Fggn0000384; cta.
DR InterPro: IPR001019; Gprotein_alpha.
DR Pfam: PF00503; G-alpha; 1.
DR PRINTS: PR00318; GPROTEIN_A.
DR ProDom: PD000281; Gprotein_alpha; 2.
DR SMART: SM00275; G-alpha; 1.
KW GTP-binding; Transducer; Multigene family; Gastrulation.
FT NP_BIND 139 146 GTP (BY SIMILARITY).
FT NP_BIND 299 303 GTP (BY SIMILARITY).
FT NP_BIND 369 372 GTP (BY SIMILARITY).
SQ SEQUENCE 457 AA; 52753 MW; D204415C4B02CEC7 CRC64;
Query Match 33.3%; Score 5; DB 1; Length 457;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 RSTPE 11
DB 103 RSTPE 107
RESULT 93
HXK4_HUMAN STANDARD; PRT; 465 AA.
ID HXK4_HUMAN
AC P35557; 005810;
DT 01-JUN-1994 (Rel. 29, Created)

DT 15-JUN-2002 (Rel. 41, last annotation update)
 DE DNA repair protein recA, chloroplast precursor.
 GN RECA OR AT1G79050 OR YUP8H12R_18 OR YUP8H12R.33.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryotes; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=94302159; PubMed=8029344;
 RA Bluet M.N., Osman M., Jagerdorf A.T.;
 RT "Genomic nucleotide sequence of a gene from Arabidopsis thaliana
 RL encoding a protein homolog of Escherichia coli RecA.";
 RN Plant Physiol. 103:673-674(1993).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016719; PubMed=11130712;
 RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 RA Dunn P., Etgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gail J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultzar L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koo H.L., Kremetska I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Malt R., Marshall A.,
 RA Maltischer J., Miranda M., Nguyen M., Nieman W.C., Osborne B.I.,
 RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Tortum M.J., Town C.D., Walker M.,
 RA Uterbeck T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 RL thaliana.";
 RN Nature 408:816-820(2000).
 [3]
 RP SEQUENCE OF 2-439 FROM N.A.
 RX MEDLINE=92390390; PubMed=1518831;
 RA Cerutti H.D., Osman M., Grandoni P., Jagerdorf A.T.;
 RT "A homolog of Escherichia coli RecA protein in plastids of higher
 RL plants.";
 RN Proc. Natl. Acad. Sci. U.S.A. 89:8068-8072(1992).
 CC -!- FUNCTION: INVOLVED IN RECOMBINATION ABILITY AND DNA STRAND
 CC TRANSFER ACTIVITY.
 CC -!- SUBCELLULAR LOCATION: Chloroplast stroma.
 CC -!- SIMILARITY: BELONGS TO THE RECA FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L15229; AAA61781.1; -;
 DR EMBL: AC0002986; AAC17051.1; -;
 DR EMBL: M98039; AAJ2855.1; -;
 DR HSSP: P26345; 1G19.
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR001553; RECA.
 DR Pfam: PF00154; recA.1.
 DR PRINTS: PR00142; RECA.1.
 DR ProDom: PD000229; RecA.1.
 DR SMART: SM00382; AAA.1.
 DR PROSITE: PS00321; RECA_1.1.
 DR PROSITE: PS50162; RECA_2.1.
 DR PROSITE: PS50163; RECA_3.1.
 KW DNA damage; DNA recombination; ATP-binding; DNA-binding; Chloroplast;

KW Transit peptide.
 FT TRANSIT 1 52 CHLOROPLAST (POTENTIAL).
 FT CHAIN 53 439 DNA REPAIR PROTEIN RECA.
 FT NP_BIND 142 149 ATP (POTENTIAL).
 SQ SEQUENCE 439 AA: 47733 MW: 35465987B37E15 CRC64:
 Query Match 33.3%; Score 5; DB 1; Length 439;
 Best Local Similarity 100.0%; Pred. No. 1; 7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 6 HRSTP 10
 Db 411 HRSTP 415
 RESULT 90
 ZP23_BRARE
 ID ZP23_BRARE STANDARD: PRT; 443 AA.
 AC P79745; P79744;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, last annotation update)
 DE POU domain protein 2P-23.
 GN POU23 OR ZP23POU OR ZP23.
 OS Brachydanio rerio (zebrafish) (Danio rerio).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 CC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Neurula;
 RX MEDLINE=97169208; PubMed=9016656;
 RA Spatiol P., Bornmann G., Gerster T.;
 RT "Class III POU genes of zebrafish are predominantly expressed in the
 RL central nervous system.";
 CC Nucleic Acids Res. 24:4874-4881(1996).
 CC -!- FUNCTION: TRANSCRIPTION FACTOR THAT MAY PLAY IMPORTANT ROLES IN
 CC PATTERNING THE EMBRYONIC BRAIN.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A
 CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN THE CENTRAL NERVOUS
 CC SYSTEM.
 CC -!- DEVELOPMENTAL STAGE: DETECTED AFTER COMPLETION OF THE GASTRULA
 CC PERIOD. MAXIMAL EXPRESSION AFTER 1 TO 2 DAYS OF DEVELOPMENT.
 CC -!- SIMILARITY: BELONGS TO THE POU TRANSCRIPTION FACTOR FAMILY.
 CC CLANS-3 SUPERFAMILY.
 CC -!- SIMILARITY: CONSTAINS 1 HOMEBOX DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Y07907; CA669215.1; -;
 DR EMBL: Y07907; CA669214.1; -;
 DR HSSP: P14859; 1OCT.
 DR ZFIN: ZDB-GENE-980526-140; pou23.
 DR InterPro: IPR001356; Homeobox.
 DR InterPro: IPR000327; POU domain.
 DR Pfam: PF00046; homeobox.1.
 DR Pfam: PF00157; pou.1.
 DR PRINTS: PR00028; POU DOMAIN.
 DR ProDom: PD000010; Homeobox.1.
 DR ProDom: PD000583; POU_domain.1.
 DR SMART: SM00389; HOX.1.
 DR SMART: SM00352; POU.1.
 DR PROSITE: PS00027; HOMEBOX_1.1.
 DR PROSITE: PS50071; HOMEBOX_2.1.

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Sytkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Massarman D.A., Weinstein G.M., Weisenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Glibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
RN [3]
RP WPM MOTIF.
RX MEDLINE-95094252; PubMed-8001118;
RA Proulx Z., Finley R.L. Jr., Kidd T., Walnwright S.M., Ingham P.W.,
RA Brent R., Ish-Horowicz D.;
RT "Groncho is required for *Drosophila* neurogenesis, segmentation, and
RT sex determination and interacts directly with hairy-related bHLH
RT proteins.";
RL Cell 79:805-815(1994).
CC -1- FUNCTION: TRANSCRIPTIONAL REPRESSOR OF GENES THAT REQUIRE A BHLH
CC PROTEIN FOR THEIR TRANSCRIPTION. LIKELY TO BE INVOLVED IN THE
CC FUNCTIONAL. RATHER THAN THE MORPHOLOGICAL DIFFERENTIATION OF
CC NEURONS. LOSS OF DPN FUNCTION RESULTS IN WEAK MOTOR ACTIVITY
CC LETHARGIC BEHAVIOR, AND DEATH. IMPLICATED IN SEX DETERMINATION AS
CC GROUCHO-DPN COMPLEX ACT DIRECTLY TO REPRESS SXL TRANSCRIPTION.
CC SUBUNIT: TRANSCRIPTION REPRESSION REQUIRES FORMATION OF A COMPLEX
CC WITH A CO-REPRESSOR PROTEIN (GROUCHO).
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN PRIMARY NEURAL PRECURSORS.
CC -1- DEVELOPMENTAL STAGE: FIRST DETECTED IN PREBLASTODERM CYCLE 12 IN
CC ALL NUCLEI. DURING MIDDLE TO LATE CYCLE 13, EXPRESSED IN EIGHT
CC STRIPES THAT OVERLAP THOSE OF THE HAIRY PROTEIN.
CC -1- DOMAIN: HAS A PARTICULAR TYPE OF BASIC DOMAIN (PRESENCE OF A
CC HELIX-INTERUPTING PROLINE) THAT BINDS TO THE N-BOX (CAGNAG),
CC RATHER THAN THE CANONICAL E-BOX (CANNTG).
CC -1- DOMAIN: THE CARBOXYL-TERMINAL WPM MOTIF IS A TRANSCRIPTIONAL
CC REPRESSION DOMAIN NECESSARY FOR THE INTERACTION WITH GROUCHO, A
CC TRANSCRIPTIONAL CO-REPRESSOR RECRUITED TO SPECIFIC TARGET DNA BY
CC HAIRY-RELATED PROTEINS.
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS. "HAIRY" SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S48025; AAB24149.1; -;
DR EMBL; AE003838; AAF59113.1; -;
DR FlyBase; FBgn0010109; dpn.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR003650; Orange.
DR Pfam; PF00010; HLH_1.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00511; ORANGE; 1.
DR PROSITE; PS00038; HLH_1; 1.
DR PROSITE; PS50888; HLH_2; 1.
DR Differentiation; Developmental protein; Nuclear protein; DNA-binding;
KW Neurogenesis; Repressor; Transcription regulation.
FT DNA_BIND 41 53 BASIC DOMAIN.
FT DOMAIN 54 98 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
FT DOMAIN 386 389 POLY-GLN.
FT DOMAIN 432 435 WPM MOTIF.
FT CONFLICT 339 339 M -> L (IN REF. 1).
SQ SEQUENCE 435 AA; 46551 MW; 9511A6C5F5019A29 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 435;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRAA 15

DB 328 ESRAA 332
|||||
RESULT 88
NTT4_HUMAN STANDARD; PRT; 439 AA.
ID NTT4_HUMAN
AC NTTH48
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Orphan sodium- and chloride-dependent neurotransmitter transporter
DE NTT4 (Fragment).
GN NTT4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Clark G.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: NOT YET KNOWN, ORPHAN TRANSPORTER.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE SODIUM-NEUROTRANSMITTER SYMPORTER
CC FAMILY (SNF).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL137790; CAC19682.1; ALT_INT.
DR InterPro; IPR000175; Na/nttran_symport.
DR Pfam; PF00209; SNF_1.
DR PRINTS; PR00176; NANEUSMPORT.
DR ProDom; PD000448; Na/nttran_symport; 2.
DR PROSITE; PS00610; NA_NEUTROTAN_SYMPT_1; PARTIAL.
DR PROSITE; PS00754; NA_NEUTROTAN_SYMPT_2; PARTIAL.
DR PROSITE; PS50267; NA_NEUTROTAN_SYMPT_3; 1.
KW Neurotransmitter transport; Transmembrane; Glycoprotein;
KW Symport.
FT NON_TER 1 1
FT TRANSMEM 17 34 5 (POTENTIAL).
FT TRANSMEM 46 67 7 (POTENTIAL).
FT DOMAIN 68 163 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 164 183 8 (POTENTIAL).
FT TRANSMEM 207 225 9 (POTENTIAL).
FT TRANSMEM 241 261 10 (POTENTIAL).
FT TRANSMEM 282 303 11 (POTENTIAL).
FT TRANSMEM 331 353 12 (POTENTIAL).
FT TRANSMEM 354 439 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 105
SQ SEQUENCE 439 AA; 49011 MW; 44634308C281C740 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 439;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 433 STES 437
8 STES 12
|||||
ID RECA_ARATH STANDARD; PRT; 439 AA.
AC Q39199; Q39200;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)

DR Pfam: PF00018; SH3; 1.
 DR PRINTS: PR00452; SH3DOMAIN.
 DR PRODOM: PD000066; SH3; 1.
 DR SMART: SM00055; FCH; 1.
 DR SMART: SM00326; SH3; 1.
 DR PROSITE: PS50133; FCH; 1.
 DR PROSITE: PS50002; SH3; 1.
 DR Endocytosis; SH3 domain; Coiled coil; Phosphorylation.
 FT DOMAIN 10 73
 FT DOMAIN 363 424
 FT DOMAIN 174 217
 FT DOMAIN 360 360
 FT CONFLICT 360 360
 FT SEQUENCE 424 AA; 48585 MW; 00475BC0321485B7 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 424;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRAA 15
 DB 128 ESRAA 132

RESULT 86
 POU1_BRARE STANDARD: PRT; 425 AA.
 AC P31366;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE POU domain protein 1 (ZFP001).
 GN POU1.
 OS Brachydanio rerio (zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OC NCBI_TaxID=7955;
 RN [1]
 RP MEDLINE=93038620; PubMed=1417821;
 RA Matsuzaki T., Amano H., Takeda H.;
 RT "A POU-domain gene of zebrafish, ZFP001, specifically expressed in
 the developing neural tissues.";
 RL Biochem. Biophys. Res. Commun. 187:1446-1453(1992).
 CC -1 SUBCELLULAR LOCATION: Nuclear.
 CC -1 TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN THE DEVELOPING
 NEURAL TISSUES.
 CC -1 SIMILARITY: BELONGS TO THE POU TRANSCRIPTION FACTOR FAMILY.
 CC -1 SIMILARITY: BELONGS TO THE POU TRANSCRIPTION FACTOR FAMILY.
 CC CLASS=3 SUBFAMILY.
 CC -1 SIMILARITY: CONTAINS 1 HOMEBOX DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation-
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: D13045; BAA02377.1; -.
 CC PIR: JH0710; JH0710.
 DR HSSP: P14859; 1OCT.
 DR ZFIN: ZDB-GENE-990415-209; POU1.
 DR InterPro: IPR001356; Homeobox.
 DR InterPro: IPR000327; POU domain.
 DR Pfam: PF00046; Homeobox; 1.
 DR Pfam: PF00157; POU; 1.
 DR PRINTS: PR00028; POUDOMAIN.
 DR PRODOM: PD000010; Homeobox; 1.
 DR PRODOM: PD000583; POU domain; 1.
 DR SMART: SM00389; HOX; 1.
 DR SMART: SM00352; POU; 1.
 DR PROSITE: PS00027; HOMEBOX_1; 1.

DR PROSITE: PS00035; POU_1; 1.
 DR PROSITE: PS00465; POU_2; 1.
 DR PROSITE: PS50071; HOMEBOX_2; 1.
 DR Homeobox; DNA-binding; Transcription regulation; Nuclear protein.
 FT DOMAIN 84 98
 FT DOMAIN 211 227
 FT DOMAIN 242 312
 FT DNA_BIND 330 389
 FT SEQUENCE 425 AA; 45626 MW; 5C50C09CE7E16D CRC64;

Query Match 33.3%; Score 5; DB 1; Length 425;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HLGPH 6
 DB 139 HLGPH 143

RESULT 87
 DPN_DROME STANDARD: PRT; 435 AA.
 AC Q26263; Q9Y384;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Deadpan protein.
 DE DPN OR CG8704.
 GN Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 RP MEDLINE=93051333; PubMed=1427077;
 RA Bier E., Vaessin H., Younger-Shepherd S., Jan L.Y., Jan Y.N.;
 RT "Deadpan, an essential pan-neural gene in Drosophila, encodes a
 helix-loop-helix protein similar to the hairy gene product.";
 RL Genes Dev. 6:2137-2151(1992).
 RN [2]
 RP MEDLINE=93051333; PubMed=10731132;
 RC STRAIN-Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktoglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borrova D., Botchan M.R., Bouck J., Brostein P., Brothier P.,
 RA Butris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cavley S., Dahlke C., Davaport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays R.A., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jajani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheller F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RX PubMed-11082044;
 RA Modregger J., Rittler B., Witter B., Paulsson M., Plomann M.;
 RT "All three PACSIN isoforms bind to endocytic proteins and inhibit
 RT endocytosis."; J. Cell Sci. 113:4511-4521(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-21100457; PubMed-11179684;
 RA Smov L., Pluvinet R., Andreu N., Estivill X., Escarceller M.;
 RT "PACSIN 3 is a novel SH3 domain cytoplasmic adapter protein of the
 RT pacsin-syndapin-FAP52 gene family."; Gene 262:199-205(2001).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE-SKIN, and uterus;
 RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 1-405 FROM N.A.
 RA Tanigami A., Fujiwara T., Ono T., Yamada K., Fujii Y., Ozaki K.,
 RA Hiroo M., Ohmori Y., Ota T., Suzuki Y., Obayashi M., Nishi T.,
 RA Shibahara T., Tanaka T., Nakamura Y., Isogai T., Sugano S.;
 RT "NED human cDNA sequencing project."; Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: May play a role in vesicle formation and transport.
 CC -1- SUBUNIT: Homo- and hetero-aggregates with other PACSINs. Binds
 CC dynamin 1, synaptotagmin, synapsin 1 and the neural Wiskott-Aldrich
 CC syndrome protein (N-WASP). Also interacts with MDC9 and MDC15 (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic. Detected at the inner aspect of
 CC the plasma membrane in cells or myotubes.
 CC -1- TISSUE SPECIFICITY: Expressed in heart and skeletal muscle,
 CC lung, kidney, adrenal gland and placenta.
 CC -1- PTM: Phosphorylated by casein kinase 2 (CK2) and protein kinase C
 CC (PKC) (Probable).
 CC -1- SIMILARITY: BELONGS TO THE PACSIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 FCH DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AF130979; AAF04472.1; -;
 DR EMBL: AF149825; AAG31023.1; -;
 DR EMBL: AF242530; AAK29207.1; -;
 DR EMBL: BC007914; AAH07914.1; -;
 DR EMBL: BC011889; AAH11889.1; -;
 DR EMBL: AK000577; BAA91267.1; -;
 DR HSSP: P29355; ISEM.
 DR Genew: HGNC:8572; PACSIN3.
 DR MIM: 606513; -;
 DR InterPro: IPR001060; Cdc15_Fes_CIP4.
 DR InterPro: IPR001452; SH3.
 DR Pfam: PF00611; FCH; 1.
 DR Pfam: PF00018; SH3; 1.
 DR PRINTS: PR00452; SH3DOMAIN.
 DR ProDom: PD000066; SH3; 1.
 DR SMART: SM00055; FCH; 1.
 DR SMART: SM00326; SH3; 1.
 DR PROSITE: PSS0133; FCH; 1.
 DR PROSITE: PSS0002; SH3; 1.
 DR Endocytosis: SH3 domain; Coiled coil; Phosphorylation.
 FT DOMAIN 10 73 FCH.
 FT DOMAIN 363 424 SH3.
 FT DOMAIN 167 217 COILED COIL (POTENTIAL).
 FT CONFLICT 61 61 A -> V (IN REF. 1).
 SQ SEQUENCE 424 AA; 48486 MW; 6DBD940AE5DF352 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 424;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 11 ESRA 15
 Db 128 ESRA 132
 RESULT 85
 PAC3_MOUSE STANDARD; PRT; 424 AA.
 AC 099JB8; Q9EQP9;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Protein kinase C and casein kinase II substrate protein 3.
 GN PACSIN3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND FUNCTION.
 RC STRAIN=C57BL/6J;
 RX MEDLINE-20534871; PubMed-11082044;
 RA Modregger J., Rittler B., Witter B., Paulsson M., Plomann M.;
 RT "All three PACSIN isoforms bind to endocytic proteins and inhibit
 RT endocytosis."; J. Cell Sci. 113:4511-4521(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-21100457; PubMed-11179684;
 RA Smov L., Pluvinet R., Andreu N., Estivill X., Escarceller M.;
 RT "PACSIN 3 is a novel SH3 domain cytoplasmic adapter protein of the
 RT pacsin-syndapin-FAP52 gene family."; Gene 262:199-205(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Neuron;
 RA Strausberg R.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: May play a role in vesicle formation and transport.
 CC -1- SUBUNIT: Homo- and hetero-aggregates with other PACSINs. Binds
 CC dynamin 1, synaptotagmin, synapsin 1 and the neural Wiskott-Aldrich
 CC syndrome protein (N-WASP). Also interacts with MDC9 and MDC15 (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic. Detected at the inner aspect of
 CC the plasma membrane in cells or myotubes (By similarity).
 CC -1- TISSUE SPECIFICITY: Highly expressed in skeletal muscle, heart and
 CC lung.
 CC -1- PTM: Phosphorylated by casein kinase 2 (CK2) and protein kinase C
 CC (PKC) (Probable).
 CC -1- SIMILARITY: BELONGS TO THE PACSIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 FCH DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AF149824; AAG31022.1; -;
 DR EMBL: AF242531; AAK29208.1; -;
 DR EMBL: BC003884; AAH03884.1; -;
 DR HSSP: P29355; ISEM.
 DR MGI: 1891410; Pacsin3.
 DR InterPro: IPR001060; Cdc15_Fes_CIP4.
 DR InterPro: IPR001452; SH3.
 DR Pfam: PF00611; FCH; 1.

```

CC CHANNEL.
CC -1- SUBMIT: HOMO- OR HETEROPOLYMERS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A (SHOWN HERE) AND B; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN BRAIN AND IMMUNE
CC SYSTEM.
CC -1- SIMILARITY: BELONGS TO THE P2X RECEPTOR FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U49395; AAB08576.1; -.
DR EMBL: U49396; AAB08577.1; -.
DR EMBL: AF016709; AAC51931.1; -.
DR EMBL: AF070573; AAC28645.1; -.
DR EMBL: AF168787; AAF43106.1; ALT_SEQ.
DR EMBL: AF168787; AAF43105.1; -.
DR GenBank: HGNC:8536; P2RX5.
DR MIM: 602836; -.
DR InterPro: IPR001429; P2X_receptor.
DR Pfam: PF00864; P2X_receptor; 1.
DR PRINTS: PRO1307; P2XRECEPTOR.
DR TIGRFAMs: TIGR00863; P2X; 1.
DR PROSITE: PS01212; P2X_RECEPTOR; 1.
DR KW Ionic channel; Transmembrane; Ion transport; Receptor; Glycoprotein;
KW Alternative splicing.
FT DOMAIN 1 30 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 31 51 1 (POTENTIAL).
FT DOMAIN 52 318 EXTRACELLULAR; CYSTEINE-RICH (POTENTIAL).
FT TRANSMEM 319 338 2 (POTENTIAL).
FT DOMAIN 340 421 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 77 77 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPIC 202 202 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 97 120 MISSING (IN ISOFORM B).
FT CONFLICT 97 98 GE -> EK (IN REF. 1).
FT CONFLICT 156 156 E -> G (IN REF. 2).
FT CONFLICT 204 204 S -> SK (IN REF. 2).
FT CONFLICT 236 237 VI -> IV (IN REF. 2).
FT CONFLICT 250 250 E -> R (IN REF. 2).
FT CONFLICT 329 329 F -> S (IN REF. 1).
FT CONFLICT 351 351 E -> O (IN REF. 1).
FT CONFLICT 398 399 KR -> NV (IN REF. 1).
SQ SEQUENCE 421 AA; 47103 MW; 6B47307AA14C2A6C CRC64;

Query Match 33.3%; Score 5; DB 1; Length 421;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 PHRST 9
Db 417 PHRST 421

RESULT 83
GAS7_RAT
ID GAS7_RAT STANDARD; PRT; 422 AA.
AC O55148;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Growth-arrest-specific protein 7 (GAS-7).
GN GAS7.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CX NCBI_TaxID=10116;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN-Sprague-Dawley; Tissue-Brain;
RA Chao C.C.-K., Kuo J.D., Su L.J., Liu C.H., Lin C.J., Cohen S.N.,
RA Lin-Chao S.;
RT "Rat gas-7: a growth-arrest-specific gene in fibroblasts is
RT preferentially expressed in terminally differentiated Purkinje neurons
RT and associated with the differentiation of cultured PC12 cells.";
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAY PLAY A ROLE IN PROMOTING MATURATION AND
CC MORPHOLOGICAL DIFFERENTIATION OF CEREBELLAR NEURONS.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: CONTAINS 1 FCH DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 MW DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ003148; CAA05907.1; ALT_INIT.
DR EMBL: AJ31902; CAA10525.1; ALT_INIT.
DR HSSP: Q13526; IPIIN.
DR InterPro: IPR001060; Cdc15_Fes_CIP4.
DR InterPro: IPR001202; MW_Rsp5_WMP.
DR Pfam: PF00397; MW; 1.
DR Pfam: PF00611; FCH; 1.
DR SMART: SM00055; FCH; 1.
DR SMART: SM00456; MW; 1.
DR PROSITE: PS01133; FCH; 1.
DR PROSITE: PS01159; MW_DOMAIN_1; 1.
DR PROSITE: PS50020; MW_DOMAIN_2; 1.
DR KW Neurogenesis; Developmental protein; Coiled coil.
FT DOMAIN 22 55 MW.
FT DOMAIN 141 233 FCH.
FT DOMAIN 254 329 COILED COIL (POTENTIAL).
SQ SEQUENCE 422 AA; 48274 MW; FD148CBADEFB01F9 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 422;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 GPHRS 8
Db 64 GPHRS 68

RESULT 84
PAC3_HUMAN
ID PAC3_HUMAN STANDARD; PRT; 424 AA.
AC Q9UKS6; Q9H331; Q9NNV9;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Protein kinase C and casein kinase substrate in neurons protein 3 (SH3
DE domain-containing protein 6511) (Endophilin I).
GN PACSIN3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Breast carcinoma;
RC MEDLINE=20002705; PubMed=10531379;
RA Howard L., Nelson K.K., Maciewicz R.A., Blobel C.P.;
RT "Interaction of the metalloprotease disintegrins MDC9 and MDC15 with
RT two SH3 domain-containing proteins, endophilin I and SH3BP1.";
RL J. Biol. Chem. 274:31693-31699(1999).
RN [2]
RP SEQUENCE FROM N.A., AND FUNCTION.

```

```

DR PROSITE: PS00344; GATA_ZN_FINGER_1; 2.
DR PROSITE: PS50114; GATA_ZN_FINGER_2; 2.
KM Transcription regulation: Activator; DNA-binding; Erythrocyte;
KM Zinc-finger; Nuclear protein; Disease mutation.
FT ZN_FING 204 228 GATA-TYPE 1.
FT ZN_FING 258 282 GATA-TYPE 2.
FT BINDING 203 203 FOG1.
FT BINDING 205 205 FOG1.
FT BINDING 208 208 FOG1.
FT BINDING 222 222 FOG1.
FT VARIANT 205 205 V -> M (IN XDAT: SEVERE IMPAIRMENT OF
FOG1 BINDING AND ERYTHROID
DIFFERENTIATION IN VITRO).
FT VARIANT 208 208 /FTID-VAR_010115.
G -> S (IN XDAT: PARTIALLY DISRUPT THE
INTERACTION WITH FOG1).
FT VARIANT 218 218 D -> G (IN XDAT: PARTIALLY DISRUPT THE
INTERACTION WITH FOG1).
FT MUTAGEN 204 204 /FTID-VAR_012707.
C->R: INCREASE OF DISSOCIATION RATE FROM
BOUND DNA.
SQ SEQUENCE 413 AA; 42751 MW; 822BD2DE14B908AD CRC64;

Query Match
Best Local Similarity 33.3%; Score 5; DB 1; Length 413;
Matches 5; Conservative 100.0%; Pred. No. 1.6e+02;

OY 8 SPES 12
DB 26 SPES 30

RESULT 81
STAL_CHICK STANDARD; PRT: 413 AA.
AC Q92182;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE CMP-N-acetylneuraminate-beta-galactoside-alpha-2,6-sialyltransferase
(EG 2.4.99.1) (Beta-galactoside alpha-2,6-sialyltransferase)
GN (Alpha 2,6-ST) (Sialyltransferase 1) (ST6Gal 1).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=94139712; PubMed=8307003;
RA Kurosawa N., Kawasaki M., Hamamoto T., Nakacka T., Lee Y.-C.,
RA Arita M., Tsuji S.;
RT "Molecular cloning and expression of chick embryo gal beta 1,4GlcNAc
alpha 2,6-sialyltransferase. Comparison with the mammalian enzyme.";
RL Eur. J. Biochem. 219:375-381(1994).
CC -I- FUNCTION: TRANSFERS SIALIC ACID FROM THE DONOR OF SUBSTRATE CMP-
SIALIC ACID TO GALACTOSE CONTAINING ACCEPTOR SUBSTRATES.
CC -I- CATALYTIC ACTIVITY: CMP-N-acetylneuraminate + beta-D-galactosyl-
1,4-acetyl-beta-D-glucosamine = CMP + alpha-N-acetylneuraminy-
2,6-beta-D-galactosyl-1,4-N-acetyl-beta-D-glucosamine.
CC -I- PATHWAY: Glycosylation
CC -I- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
FORN IN TRANS CISTERNAE OF GOLGI, SOLUBLE FORM IN BODY FLUIDS.
CC -I- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN TESTES, BRAIN AND LIVER
AND TO A LESSER EXTENT IN LUNG AND HEART.
CC -I- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
PROTEOLYTIC PROCESSING.
CC -I- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 29.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb.ch).
CC -----
CC EMBL: X75558; CAA53235.1; -.
DR InterPro: IPR001675; GT_29.
DR Pfam: PF00777; Glyco_transf_29; 1.
KM Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
KM Signal-anchor; Golgi stack.
FT DOMAIN 1 9 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 10 26 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
(POTENTIAL).
FT DOMAIN 27 413 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 129 129 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 155 155 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 167 167 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 413 AA; 47392 MW; 19B8CCD361ED137D CRC64;

Query Match
Best Local Similarity 33.3%; Score 5; DB 1; Length 413;
Matches 5; Conservative 100.0%; Pred. No. 1.6e+02;

OY 1 SHLG 5
DB 202 SHLG 206

RESULT 82
P2X5_HUMAN STANDARD; PRT: 421 AA.
AC Q93086; Q93087; Q43450; Q75540; Q9NZV0;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE P2X purinoceptor 5 (ATP receptor) (P2X5) (Purinergic receptor).
GN P2X5 OR P2X5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS A AND B).
RA Tokuyama Y., Meret L., Chen X., Rouard M., Bell G.I.;
RT "Cloning of human p2x purinoceptor new subtype (P2X5).";
RL Submitted (SEP-1996) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM A).
RC TISSUE=Brain;
RX MEDLINE=98074936; PubMed=9414125;
RA Le K.-T., Paquet M., Nouel D., Babiniski K., Seguela P.;
RT "Primary structure and expression of a naturally truncated human P2X
ATP receptor subunit from brain and immune system.";
RL FEBS Lett. 418:195-199(1997).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM B).
RC TISSUE=Brain;
RX Yu W., Gibbs R.A.;
RL Submitted (JUN-1998) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A. (ISOFORMS A AND B).
RX MEDLINE=20138496; PubMed=10673375;
RA Touchman J.W., Aikster Y., Dietrich N.L., Maduro V.V., McDowell G.,
RA Shchelersuk V., Bouffard G.G., Beckstrom-Sternberg S.M., Gahl W.A.,
RA Green E.D.;
RT "The genomic region encompassing the nephropathic cystinosis gene
(CTNS): complete sequencing of a 200-kb segment and discovery of a
novel gene within the common cystinosis-causing deletion.";
RL Genome Res. 10:165-173(2000).
CC -I- FUNCTION: RECEPTOR FOR ATP THAT ACTS AS A LIGAND GATED ION

```

RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
 RA Waterston R., Watson A., Wainstock L., Wilkinson-Sprat J.,
 RA Woldman P.,
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RL Nature 368:32-38(1994).
 RP REVISIONS.
 RA Waterston R.;
 RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases:
 CC -1- CATALYTIC ACTIVITY: 3-methylbutanoyl-CoA + ETF -> 3-methylbut-2-
 CC enoyl-CoA + reduced ETF.
 CC -1- COFACTOR: FAD.
 CC -1- PATHWAY: Leucine catabolism.
 CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE ACYL-CoA DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L16622; AAA27913.2; -;
 DR PIR: S44743; S44743.
 DR HSSP: P26440; 1YH.
 DR Wompep; C02D5.1; CE24778.
 DR InterPro: IPR001552; Acyl-CoA_dh.
 DR Pfam: PF00441; Acyl-CoA_dh. 1.
 DR Pfam: PF02770; Acyl-CoA_dh. 1.
 DR Pfam: PF02771; Acyl-CoA_dh. N: 1.
 DR PROSITE: PS00072; ACYL-CoA_DH.1: 1.
 DR PROSITE: PS00073; ACYL-CoA_DH.2: 1.
 DR Hypothetical protein; Oxidoreductase; Flavoprotein; FAD.
 KW AC_SITE 265 265 BY SIMILARITY.
 FT SEQUENCE 408 AA; 45326 MW; A6351C6DCDCA962D CRC64;
 SQ
 Query March 33.3%; Score 5; DB 1; Length 408;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RX MEDLINE-90114418; PubMed-2104960;
 RA Trautnor C.D., Evans T., Felsenfeld G., Boguski M.S.;
 RT "Structure and evolution of a human erythroid transcription factor.";
 RL Nature 343:92-96(1990).
 RP SEQUENCE FROM N.A.
 RA Blechschmidt K., Nvakatura G., Strom T.M., Drescher B., Menzel U.,
 RA Meindl A., Rosenthal A.; the EMBL/GenBank/DBJ databases.
 RL Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP BINDING TO FOG1, VARIANT XDAT MET-205, AND MUTAGENESIS OF CYS-204.
 RP TISSUE-Peripheral blood;
 RX MEDLINE-20164324; PubMed-10700180;
 RA Nichols K.E., Crispino J.D., Poncz M., White J.G., Orkin S.H.,
 RA Maris J.M., Weiss M.J.;
 RT "Familial dyserythropoietic anaemia and thrombocytopenia due to an
 RT inherited mutation in GATA1.";
 RL Nat. Genet. 24:266-270(2000).
 RP [5]
 RP VARIANT XDAT GLY-218.
 RX PubMed-1118466;
 RA Frieson K., Devriendt K., Matthijs G., Van Hoof A., De Vos R., Thys C.,
 RA Minner K., Hoylaerts M.F., Vermeylen J., Van Geet C.;
 RT "Platelet characteristics in patients with x-linked
 RT macrothrombocytopenia because of a novel GATA1 mutation.";
 RL Blood 98:85-92(2001).
 RP [6]
 RP VARIANT XDAT SER-208.
 RX MEDLINE-21531306; PubMed-11675338;
 RA Menafey M.G., Newton A.L., Gandhi M.J., Crossley M., Drachman J.G.;
 RT "X-linked thrombocytopenia caused by a novel mutation of GATA-1.";
 RL Blood 98:2681-2688(2001).
 CC -1- FUNCTION: TRANSCRIPTIONAL ACTIVATOR WHICH PROBABLY SERVES AS A
 CC GENERAL SWITCH FACTOR FOR ERYTHROID DEVELOPMENT. IT BINDS TO DNA
 CC SITES WITH THE CONSENSUS SEQUENCE (A/T)GATA(A/G) WITHIN REGULATORY
 CC REGIONS OF GLOBIN GENES AND OF OTHER GENES EXPRESSED IN ERYTHROID
 CC CELLS.
 CC -1- SUBUNIT: THE AMINO-TERMINAL ZINC FINGER INTERACTS WITH FOG1.
 CC -1- TISSUE SPECIFICITY: ERYTHROCYTE.
 CC -1- DOMAIN: THE TWO FINGERS ARE FUNCTIONALLY DISTINCT AND COOPERATE TO
 CC ACHIEVE SPECIFIC, STABLE DNA BINDING. THE FIRST FINGER IS
 CC NECESSARY ONLY FOR FULL SPECIFICITY AND STABILITY OF BINDING,
 CC WHEREAS THE SECOND ONE IS REQUIRED FOR BINDING (BY SIMILARITY).
 CC -1- DISEASE: DEFECTS IN GATA1 ARE THE CAUSE OF X-LINKED
 CC DYSERYTHROPOIETIC ANEMIA AND THROMBOCYTOPENIA (XDAT). A DISORDER
 CC CHARACTERIZED BY ERYTHROCYTES WITH ABNORMAL SIZE AND SHAPE, AND
 CC PAUCITY OF PLATELETS IN PERIPHERAL BLOOD. THE BONE MARROW CONTAINS
 CC ABUNDANT AND ABNORMALLY SMALL MEGAKARYOCYTES.
 CC -1- SIMILARITY: CONTAINS 2 GATA-TYPE ZINC FINGERS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M30601; AAA35885.1; -;
 DR EMBL: X17254; CAA35120.1; -;
 DR EMBL: AF196971; AAF06806.1; -;
 DR PIR: A34888; A34888.
 DR PIR: S07121; S07121.
 DR HSSP: P17679; 1GNF.
 DR TRANSFAC: T00306; -;
 DR GeneW; HGNC:4170; GATA1.
 DR MIM: 305371; -;
 DR InterPro: IPR000679; znf_GATA.
 DR InterPro: IPR001164; hrip_1like.
 DR Pfam: PF00320; GATA. 2.
 DR PRINTS: PR00619; GATAZNFINGER.
 DR SMART: SM00401; znf_GATA; 2.

```

DR Pfam: PF00614; PLDC; 2.
DR SMART: SM00155; PLDC; 2.
DR PROSITE: PS00035; PLD; 1.
KW Transmembrane; Late protein.
FT TRANSMEM 17 35
FT DOMAIN 133 153
FT PLD PHOSPHODIESTERASE.
SQ SEQUENCE 388 AA; 42660 MW; 089A697B327AA5D6 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPE 11
Db 58 RSTPE 62

RESULT 77
CENC_SHEEP
ID CENC_SHEEP STANDARD; PRT; 402 AA.
AC P49453;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Centromere protein C (CENP-C) (Centromere autoantigen C) (Fragment).
GN CENP-C.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA MEDLINE=97049078; PubMed=8893808;
RA Burkin D.J., Jones C.A., Burkin H.R., McGrew J.A., Broad T.E.;
RA "Sheep CENP-B and CENP-C genes show a high level of sequence similarity
RT and conserved synteny with their human homologs.";
RL Cytogenet. Cell Genet. 74:86-89(1996).
CC -!- FUNCTION: COMPONENT OF THE INNER KINETOCORE PLATE. REQUIRED FOR
CC NORAL KINETOCORE ASSEMBLY.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: U35657; AAT9099.1; -
KW Chromosomal protein; Nuclear protein; DNA-binding; Phosphorylation;
KW Centromere.
FT NON_TER 1
SQ SEQUENCE 402 AA; 45065 MW; 9EA59F16C3EA89AB CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STPE 12
Db 177 STPE 181

RESULT 78
CPEF_STRGO
ID CPEF_STRGO STANDARD; PRT; 402 AA.
AC P18327;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)

```

```

DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Cytochrome P450-SU2 (EC 1.14.1.1) (P450-CYB1) (CYP105B1).
GN CYP105B1 OR SUBC.
OS Streptomyces griseolus.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1909;
RN [1]
RP SEQUENCE FROM N.A. AND SEQUENCE OF 1-32.
RC STRAIN=ATCC 11796;
RX MEDLINE=90264332; PubMed=2345149;
RA Omer C.A., Lenstra R., Little P.J., Dean C., Tepperman J.M.,
RA Leto K.J., Romesser J.A., O'Keefe D.P.;
RT "Genes for two herbicide-inducible cytochromes P-450 from
RT Streptomyces griseolus."
RL J. Bacteriol. 172:3335-3345(1990).
CC -!- FUNCTION: METABOLISM OF A NUMBER OF SULFONYUREA HERBICIDES.
CC -!- INDUCTION: BY HERBICIDES.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M32239; AAA6825.1; -
DR PIR: B35401; B35401.
DR HSSP: P23295; 1CNM.
DR InterPro: IPR001128; Cytochrome_P450.
DR Pfam: PF00067; P450; 1.
DR PROSITE: PS00086; CYTOCHROME_P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Heme.
FT INIT_MET 0
FT BINDING 351 351
SQ SEQUENCE 402 AA; 44278 MW; E3B67F6F7C26D9D CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPE 11
Db 183 RSTPE 187

RESULT 79
IVD_CAEEL
ID IVD_CAEEL STANDARD; PRT; 408 AA.
AC P34275;
DT 01-FEB-1994 (Rel. 28, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable isovaleryl-CoA dehydrogenase (EC 1.3.99.10) (IVD).
GN C02D5.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Alnsough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones K., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkhen R.,
RA Sims M., Smailon N., Smith A., Smith M., Sonhammer E., Staden R.,

```

OS Paracoccus denitrificans.
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Paracoccus.
OX NCBI_TaxID=266;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-Pd 1222;
RA MEDLINE=93324575; PubMed=7601147;
RA der Palen C.J., Slotboom D.J., Jongejan L., Reijnders W.N.,
Harms N., Duine J.A., van Spanning R.J.M.;
RT "Metabolic analysis of mau genes involved in methylamine metabolism
in Paracoccus denitrificans.";
RL Eur. J. Biochem. 230:860-871(1995).
CC -1- FUNCTION: INVOLVED IN METHYLAMINE METABOLISM. ESSENTIAL FOR THE
CC MATURATION OF THE BETA SUBUNIT OF MADH, PRESUMABLY VIA A STEP IN
CC THE BIOSYNTHESIS OF TRYPTOPHAN TRYPTOPHYLDIOLINONE (TTO), THE
CC COFACTOR OF MADH.
CC -1- PATHWAY: Methylamine utilization.
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- PTM: BINDS 2 HEMES (POTENTIAL).
CC -1- SIMILARITY: SOME, TO P.AERUGINOSA CYTOCHROME C PEROXIDASE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; U15028; AAA86467.1; -.
DR HSSP; P14532; IE87.
DR InterPro; IPR004852; CYTC_Maug.
DR InterPro; IPR000345; CYTC_heme_bind.
DR Pfam; PF03150; CCP_Maug.1.
DR PROSITE; PS00190; CYTOCHROME_C; 2.
KW Heme; Electron transport; periplasmic; signal.
FT SIGNAL 1 20
FT CHAIN 1 387
FT BINDING 51 51 HEME 1 (COVALENT) (BY SIMILARITY).
FT BINDING 54 54 HEME 1 (COVALENT) (BY SIMILARITY).
FT METAL 55 55 IRON 1 (HEME PROXIMAL) (BY SIMILARITY).
FT BINDING 221 221 HEME 2 (COVALENT) (BY SIMILARITY).
FT BINDING 224 224 HEME 2 (COVALENT) (BY SIMILARITY).
FT METAL 225 225 IRON 2 (HEME PROXIMAL) (BY SIMILARITY).
SQ SEQUENCE 387 AA; 42230 MW; DDC9618235D6838E CRC64;
Query Match 33.3%; Score 5; DB 1; Length 387;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 11 ESRRA 15
Db 380 ESRRA 384
RESULT 75
VENV_MCV1
ID VENV_MCV1 STANDARD; PRT; 388 AA.
AC P26579;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Major envelope protein (43 kDa protein) (p43K).
GN P43K.
OS Molluscum contagiosum virus subtype 1 (MCV1).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Molluscipoxvirus.
OX NCBI_TaxID=10280;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91251209; PubMed=2041084;
RA Blake N.W., Porter C.D., Archard L.C.;

RT "Characterization of a molluscum contagiosum virus homolog of the
RT vaccinia virus p37K major envelope antigen.";
RL J. Virol. 65:3583-3589(1991).
CC -1- FUNCTION: THIS PROTEIN IS THE MAJOR ANTIGEN ON THE ENVELOPE OF
CC EXTRACELLULAR VIRUS.
CC -1- SIMILARITY: TO VACCINIA VIRUS 37 kDa ENVELOPE PROTEIN.
CC -1- SIMILARITY: CONTAINS 1 PLD PHOSPHODIESTERASE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M63486; AAA46548.1; -.
DR PIR; A40340; WMVZM1.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS50035; PLD; 1.
KW Transmembrane; Late protein.
FT TRANSMEM 17 35
FT TRANSMEM 133 153
FT DOMAIN 310 337
SQ SEQUENCE 388 AA; 42862 MW; 54697FA9DA796A6F CRC64;
Query Match 33.3%; Score 5; DB 1; Length 388;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 7 RSTPE 11
Db 58 RSTPE 62
RESULT 76
VENV_MCV2
ID VENV_MCV2 STANDARD; PRT; 388 AA.
AC P25392;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Major envelope protein (43 kDa protein) (p43K).
GN P43K.
OS Molluscum contagiosum virus subtype 2 (MCV2).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Molluscipoxvirus.
OX NCBI_TaxID=10281;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91251209; PubMed=2041084;
RA Blake N.W., Porter C.D., Archard L.C.;"Characterization of a molluscum contagiosum virus homolog of the
RT vaccinia virus p37K major envelope antigen.";
RL J. Virol. 65:3583-3589(1991).
CC -1- FUNCTION: THIS PROTEIN IS THE MAJOR ANTIGEN ON THE ENVELOPE OF
CC EXTRACELLULAR VIRUS.
CC -1- SIMILARITY: TO VACCINIA VIRUS 37 kDa ENVELOPE PROTEIN.
CC -1- SIMILARITY: CONTAINS 1 PLD PHOSPHODIESTERASE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M63487; AAA46549.1; -.
DR PIR; B40340; WMVZU2.
DR InterPro; IPR001736; PLD.

Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 GPHRS 8
|||||

Db 268 GPHRS 272

RESULT 72

FOS_MOUSE STANDARD; PRT; 380 AA.

AC P01101; 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 01, Last annotation update)
DE Proto-oncogene protein c-fos (cellular oncogene fos).
GN FOS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83180421; PubMed=6301687;
RA Van Beveren C., Van Straaten F., Curran T., Mueller R., Verma I.M.;
RT "Analysis of FBJ-MuSV provirus and c-fos (mouse) gene reveals that
RT viral and cellular fos gene products have different carboxy
RT terminl.";
RL Cell 32:1241-1255(1983).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85270451; PubMed=2991903;
RA Mellink F., Curran T., Miller A.D., Verma I.M.;
RT "Removal of a 67-base-pair sequence in the noncoding region of
RT protooncogene fos converts it to a transforming gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:4987-4991(1985).
CC -1- FUNCTION: NUCLEAR PHOSPHOPROTEIN WHICH FORMS A TIGHT BUT NON-
CC FACTOR. C-FOS HAS A CRITICAL FUNCTION IN REGULATING THE
CC DEVELOPMENT OF CELLS DESTINED TO FORM AND MAINTAIN THE SKELETON.
CC IT IS THOUGHT TO HAVE AN IMPORTANT ROLE IN SIGNAL TRANSDUCTION,
CC CELL PROLIFERATION AND DIFFERENTIATION.
CC -1- SUBUNIT: HETERODIMER.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- INDUCTION: C-FOS EXPRESSION INCREASES UPON A VARIETY OF STIMULI,
CC INCLUDING GROWTH FACTORS, CYTOKINES, NEUROTRANSMITTERS,
CC POLYPEPTIDE HORMONES, STRESS AND CELL INJURY.
CC -1- SIMILARITY: BELONGS TO THE BZIP FAMILY. FOS SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: V00727; CAA24105.1; -
DR EMBL: J00370; AAA96699.1; -
DR PIR: A01343; TVMSF.
DR HSSP: P01100; 1FOS.
DR TRANSFAC: T00122; -
DR MGD: MGI:95574; Fos.
DR InterPro: IPR000837; Leuzip_Fos.
DR InterPro: IPR004827; TF_bZIP.
DR Pfam: PF00170; bZIP; 1.
DR PRINTS: PR00042; LEUZIPRPOS.
DR SMART: SM00338; BRLZ; 1.
DR PROSITE: PS00036; BZIP_BASIC; 1.
DR Proto-oncogene: Nuclear protein; Phosphorylation; DNA-binding.
KW DNA_BIND 139 160 BASIC MOTIF.
FT DOMAIN 165 193 LEUCINE-ZIPPER.
SQ SEQUENCE 380 AA; 40838 MW; 475966265952B624 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STRES 12
|||||

Db 231 STRES 235

RESULT 73

FOS_MSVEB STANDARD; PRT; 381 AA.

AC P01102; 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE p55-v-fos transforming protein.
GN V-FOS.
OS FBJ murine osteosarcoma virus.
OC Viruses; Retroid viruses; Retroviridae; Mammalian type C retroviruses.
OX NCBI_TaxID=11805;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83180421; PubMed=6301687;
RA Van Beveren C., Van Straaten F., Curran T., Mueller R., Verma I.M.;
RT "Analysis of FBJ-MuSV provirus and c-fos (mouse) gene reveals that
RT viral and cellular fos gene products have different carboxy
RT terminl.";
RL Cell 32:1241-1255(1983).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85270451; PubMed=2991903;
RA Mellink F., Curran T., Miller A.D., Verma I.M.;
RT "Removal of a 67-base-pair sequence in the noncoding region of
RT protooncogene fos converts it to a transforming gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:4987-4991(1985).
CC -1- FUNCTION: NUCLEAR PHOSPHOPROTEIN WHICH FORMS A TIGHT BUT NON-
CC FACTOR. C-FOS HAS A CRITICAL FUNCTION IN REGULATING THE
CC DEVELOPMENT OF CELLS DESTINED TO FORM AND MAINTAIN THE SKELETON.
CC IT IS THOUGHT TO HAVE AN IMPORTANT ROLE IN SIGNAL TRANSDUCTION,
CC CELL PROLIFERATION AND DIFFERENTIATION.
CC -1- SUBUNIT: HETERODIMER.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- INDUCTION: C-FOS EXPRESSION INCREASES UPON A VARIETY OF STIMULI,
CC INCLUDING GROWTH FACTORS, CYTOKINES, NEUROTRANSMITTERS,
CC POLYPEPTIDE HORMONES, STRESS AND CELL INJURY.
CC -1- SIMILARITY: BELONGS TO THE BZIP FAMILY. FOS SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: V01184; CAA24505.1; -
DR PIR: A01344; TVMVJ.
DR HSSP: P01100; 1FOS.
DR TRANSFAC: T01452; -
DR InterPro: IPR000837; Leuzip_Fos.
DR InterPro: IPR004827; TF_bZIP.
DR Pfam: PF00170; bZIP; 1.
DR PRINTS: PR00042; BRLZ; 1.
DR SMART: SM00338; BRLZ; 1.
DR PROSITE: PS00036; BZIP_BASIC; 1.
DR Oncogene: Nuclear protein; Phosphorylation; DNA-binding.
KW DNA_BIND 139 159 BASIC MOTIF.
FT DOMAIN 165 193 LEUCINE-ZIPPER.
SQ SEQUENCE 381 AA; 41647 MW; D15FDB28949A9D CRC64;

Query Match 33.3%; Score 5; DB 1; Length 381;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STRES 12
|||||

Db 231 STRES 235

RESULT 74

MAUG_PARDE STANDARD; PRT; 387 AA.

AC Q51658; 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Methylamine utilization protein maug precursor.
GN MAUG.

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL591782; CAC41575.1; -
DR EMBL: AF008187; AAB70169.2; -
DR InterPro: IPR001238; Recf.
DR InterPro: IPR003395; SMC_N.
DR Pfam: PF02463; SMC_N; 1.
DR TIGRfams: TIGR00611; recf; 1.
DR PROSITE: PS00617; Recf_1; 1.
DR PROSITE: PS00618; Recf_2; 1.
KW DNA damage: DNA replication; DNA-binding; SOS response; DNA repair;
KW ATP-binding; Complete proteome.
FT NP_BIND 34 41 ATP (POTENTIAL).
FT CONFLICT 169 176 RNRLSER -> ATGFSRNS (IN REF. 2).
SQ SEQUENCE 374 AA; 40596 MW; E81F5B937ED05B55 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 374;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 GPHRS 8
DB 267 GPHRS 271

RESULT 70
PANE_YEAST STANDARD; PRT; 379 AA.
ID PANE_YEAST
AC P38787;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 2-dehydropancoate 2-reductase (EC 1.1.1.169) (ketopantoate reductase)
GN (KPR)
OS PANS OR YHR063C.
OC Saccharomycetes cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=5286; / AB972;
RX MEDLINE=94378003; PubMed=8091229;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
RA Du Z., Faveille A., Fulton L., Gattung S., Geisel C., Kirsten J.,
RA Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,
RA Latreille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,
RA Nhan M., Rifkin L., Riles L., St Peter H., Trevisan E., Vaughan K.,
RA Vignati D., Wilcox L., Woldman P., Waterston R., Wilson R.,
RA Vaudin M.;
RT "Complete nucleotide sequence of *Saccharomyces cerevisiae* chromosome
VIII.";
RL Science 265:2077-2082(1994).
CC -1- FUNCTION: Catalyzes the NADPH-dependent reduction of ketopantoate
CC into panthoic acid (By similarity).
CC -1- CATALYTIC ACTIVITY: (R)-panthoate + NADP(+) = 2-dehydropancoate +
CC NADPH.
CC -1- PATHWAY: Pantothenate biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE KETOPANTOATE REDUCTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U00061; AAB68390.1; -
DR PIR: S46711; S46711.
DR SGD: S0001105; YHR063C.

DR InterPro: IPR003710; Apba.
DR Pfam: PF02538; Apba; 1.
KW Pyrimidine biosynthesis; Pantothenate biosynthesis; Oxidoreductase;
KW NADP.
FT NP_BIND 13 18 NADP (POTENTIAL).
FT ACT_SITE 224 224 BY SIMILARITY.
FT ACT_SITE 328 328 BY SIMILARITY.
SQ SEQUENCE 379 AA; 42821 MW; 22BCAB5652AE1A93 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 PRST 9
DB 4 PRST 8

RESULT 71
REFC_RHILO STANDARD; PRT; 379 AA.
ID RECF_RHILO
AC Q98BH1;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA replication and repair protein recf.
GN RECF OR MLI5578.
OS *Rhizobium loti* (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT *Mesorhizobium loti*.";
RL DNA Res. 7:331-338(2000).
CC -1- FUNCTION: The recf protein is involved in DNA metabolism. It is
CC required for DNA replication and normal SOS inducibility. Recf
CC binds preferentially to single-stranded, linear DNA. It also seems
CC to bind ATP (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE RECF FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP003007; BAB52001.1; -
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR01238; Recf.
DR InterPro: IPR003395; SMC_N.
DR Pfam: PF02463; SMC_N; 1.
DR SMART: SM00382; AAA; 1.
DR TIGRfams: TIGR00611; recf; 1.
DR PROSITE: PS00617; Recf_1; 1.
DR PROSITE: PS00618; Recf_2; 1.
KW DNA damage: DNA replication; DNA-binding; SOS response; DNA repair;
KW ATP-binding; Complete proteome.
FT NP_BIND 34 41 ATP (POTENTIAL).
FT CONFLICT 34 41 OC0BDD4EC4F5BECF CRC64;
SQ SEQUENCE 379 AA; 40548 MW; 0C0BDD4EC4F5BECF CRC64;

Query Match 33.3%; Score 5; DB 1; Length 379;

```

FT TRANSEM 66 91 2 (POTENTIAL).
FT DOMAIN 92 111 EXTRACELLULAR (POTENTIAL).
FT TRANSEM 112 132 3 (POTENTIAL).
FT DOMAIN 133 151 CYTOPLASMIC (POTENTIAL).
FT TRANSEM 152 176 4 (POTENTIAL).
FT DOMAIN 177 198 EXTRACELLULAR (POTENTIAL).
FT TRANSEM 199 223 5 (POTENTIAL).
FT DOMAIN 224 262 CYTOPLASMIC (POTENTIAL).
FT TRANSEM 263 286 6 (POTENTIAL).
FT DOMAIN 287 299 EXTRACELLULAR (POTENTIAL).
FT TRANSEM 300 323 7 (POTENTIAL).
FT DOMAIN 324 358 CYTOPLASMIC (POTENTIAL).
FT DISULFID 109 187 BY SIMILARITY.
FT CARBOHYD 3 3 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 6 6 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 150 150 A -> R (IN REF. 2).
SQ SEQUENCE 358 AA; 39675 MW; 1F6786C3B3258AEC CRC64;

```

```

Query Match 33.3%; Score 5; DB 1; Length 358;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 10 PESRA 14
    |||||
Db 104 PESRA 108

```

```

RESULT 68
ID HMH2_DUGT1 STANDARD; PRT; 363 AA.
AC 000401;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Homeobox protein DTH-2.
GN DTH-2.
OS Dugesia tigrina (Planarian).
OC Eukaryota; Metazoa; Platyhelminthes; Turbellaria; Seriata; Tricladida;
OC Paludicola; Dugesillidae; Glirardia.
OX NCBI_Taxid=6162;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE-93387216; PubMed-8104142;
RA Garcia-Fernandez J., Baguna J., Salo E.;
RT "Genomic organization and expression of the planarian homeobox genes
   Dth-1 and Dth-2.";
RL Development 118:241-253(1993).
RN 12
RP SEQUENCE FROM N.A.
RX MEDLINE-91334461; PubMed-1714599;
RA Salo E., Garcia-Fernandez J., Baguna J.;
RT "Planarian homeobox genes: cloning, sequence analysis, and
   expression.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:7338-7342(1991).
CC -1- FUNCTION: THIS PROTEIN MIGHT BE INVOLVED IN DETERMINATION AND/OR
CC DIFFERENTIATION OF NERVE CELLS IN THE CONTINUOUS REPLACEMENT OF
CC NEURONS IN THE CEPHALIC REGION.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- TISSUE SPECIFICITY: INTERSTINE AND UNIDENTIFIED PERIPHERAL
CC PARANCHYMAL CELLS. SLIGHTLY HIGHER LEVELS IN THE CEPHALIC REGION
CC COMPARED TO OTHER BODY REGIONS.
CC -1- SIMILARITY: BELONGS TO THE NK-2 HOMEBOX FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL: X69202; CAA49140.1; -
DR EMBL: X56500; CAA39855.1; -
DR PIR: B41151; B41151.
DR PIR: S33702; S33702.
DR HSSP: P23441; 1FTF.
DR InterPro: IPR001356; Homeobox.
DR Pfam: PF00046; homeobox.1.
DR PRINTS: PR00024; HOMEBOX.
DR PRODOM: PD000010; Homeobox; 1.
DR SMART: SM00389; HOX; 1.
DR PROSITE: PS00027; HOMEBOX_1; 1.
DR PROSITE: PS00711; HOMEBOX_2; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein.
FT DOMAIN 29 42 SPR-RICH
FT DOMAIN 62 94 ASN/SER-RICH.
FT DNA_BIND 133 192 HOMEBOX.
FT DOMAIN 319 331 SER-RICH.
FT DOMAIN 348 363 ASN/SER-RICH.
SQ SEQUENCE 363 AA; 40338 MW; 4FDEDE8E60CE2C1 CRC64;

```

```

Query Match 33.3%; Score 5; DB 1; Length 363;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 7 RSRPE 11
    |||||
Db 231 RSRPE 235

```

```

RESULT 69
ID RECF_RHME STANDARD; PRT; 374 AA.
AC P56903; O30497;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA replication and repair protein recf.
GN RECF OR R00188 OR SMC02863.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_Taxid=382;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE-21396507; PubMed-11481430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masny D.,
RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ransperger U.,
RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
   Sinorhizobium meliloti strain 1021.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
RN 12
RP SEQUENCE OF 1-176 FROM N.A.
RC STRAIN=SU47 / 1021;
RX MEDLINE-98361913; PubMed-9696772;
RA Bardin S.D., Voegelé R.F., Finan T.M.;
RT "Phosphate assimilation in Rhizobium (Sinorhizobium) meliloti:
   Identification of a pit-1-like gene.";
RL J. Bacteriol. 180:4219-4226(1998).
CC -1- FUNCTION: THE RECF PROTEIN IS INVOLVED IN DNA METABOLISM: IT IS
CC BINDS PREFERENTIALLY TO SINGLE-STRANDED, LINEAR DNA. IT ALSO SEEMS
CC TO BIND ANP (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- SIMILARITY: BELONGS TO THE RECF FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

ID PE22_RAT STANDARD; PRT; 357 AA.
 AC Q62928;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Prostaglandin E2 receptor, EP2 subtype (Prostanoid EP2 receptor) (PGE
 DE receptor, EP2 subtype).
 GN PTGER2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RA MEDLINE=98103308; PubMed=9440134;
 RA Nemoto K., Pilbeam C.C., Blak S.R., Reisz L.G.;
 RT "Molecular cloning and expression of a rat prostaglandin E2 receptor
 RT of the EP2 subtype.";
 RL Prostaglandins 54:713-725(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Spleen;
 RX MEDLINE=98141044; PubMed=9537820;
 RA Bole V., Stocco R., Sawyer N., Slipek D.M., Ungria M.D.,
 RA Neuschaefer-Rube F., Puschel G.P., Metters K.M., Abramovitz M.;
 RT "Molecular cloning and characterization of the four rat prostaglandin
 RT E2 prostanoid receptor subtypes.";
 RL Eur. J. Pharmacol. 340:227-241(1997).
 CC -1- FUNCTION: RECEPTOR FOR PROSTAGLANDIN E2 (PGE2). THE ACTIVITY OF
 CC THIS RECEPTOR IS MEDIATED BY G-S PROTEINS THAT STIMULATES
 CC ADENYLATE CYCLASE. THE SUBSEQUENT RAISE IN INTRACELLULAR CAMP IS
 CC RESPONSIBLE FOR THE RELAXING EFFECT OF THIS RECEPTOR ON SMOOTH
 CC MUSCLE (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U48858; AA97889.1; -;
 DR EMBL; U94708; AAB53325.1; -;
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PROSITE: PS00237; G-PROTEIN_RECPT_F1_1; 1.
 DR PROSITE: PS0262; G-PROTEIN_RECPT_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein.
 FT DOMAIN 1 24
 FT TRANSMEM 25 48
 FT DOMAIN 49 66
 FT TRANSMEM 67 92
 FT DOMAIN 93 112
 FT TRANSMEM 113 133
 FT DOMAIN 134 152
 FT TRANSMEM 153 177
 FT DOMAIN 178 199
 FT TRANSMEM 200 224
 FT TRANSMEM 225 262
 FT TRANSMEM 263 286
 FT TRANSMEM 287 299
 FT TRANSMEM 300 323
 FT DOMAIN 324 357
 FT DISULFID 110 188
 FT CARBOHYD 6
 FT SEQUENCE 357 AA; 39771 MW; AAB22AB280AEEFB0 CRC64;
 SQ
 Query Match 33.3%; Score 5; DB 1; Length 357;
 Best Local Similarity 100.0%; Pred. No. 1; 4e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 10 PESRA 14
 |||||
 Db 105 PESRA 109
 RESULT 67
 ID PE22_HUMAN STANDARD; PRT; 358 AA.
 AC P43116;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Prostaglandin E2 receptor, EP2 subtype (Prostanoid EP2 receptor) (PGE
 DE receptor, EP2 subtype).
 GN PTGER2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=94359483; PubMed=8078484;
 RA Regan J.W., Bailey T.J., Pepperl D.J., Pierce K.L., Bogardus A.M.,
 RA Donello J.E., Fairbairn C.E., Kedzie K.M., Woodward D.F., Gill D.W.;
 RT "Cloning of a novel human prostaglandin receptor with characteristics
 RT of the pharmacologically defined EP2 subtype.";
 RL Mol. Pharmacol. 46:213-220(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Oakley C.J.;
 RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99453300; PubMed=10521663;
 RA Smock S.L., Pan L.C., Castleberry T.A., Lu B., Mather R.J.,
 RA Owen T.A.;
 RT "Cloning, structural characterization, and chromosomal localization of
 RT the gene encoding the human prostaglandin E2 receptor EP2 subtype.";
 RL Gene 237:393-402(1999).
 CC -1- FUNCTION: RECEPTOR FOR PROSTAGLANDIN E2 (PGE2). THE ACTIVITY OF
 CC THIS RECEPTOR IS MEDIATED BY G-S PROTEINS THAT STIMULATES
 CC ADENYLATE CYCLASE. THE SUBSEQUENT RAISE IN INTRACELLULAR CAMP IS
 CC RESPONSIBLE FOR THE RELAXING EFFECT OF THIS RECEPTOR ON SMOOTH
 CC MUSCLE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: PLACENTA AND LUNG.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U19487; AA61681.1; -;
 DR EMBL; X83868; CA58749.1; -;
 DR EMBL; AF134202; AAD4177.1; -;
 DR EMBL; AF134201; AAD4177.1; JOINED.
 DR GeneW: HGNC:9594; PTGER2.
 DR MIM: 176804; -;
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PROSITE: PS00237; G-PROTEIN_RECPT_F1_1; 1.
 DR PROSITE: PS0262; G-PROTEIN_RECPT_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein.
 FT DOMAIN 1 23
 FT TRANSMEM 24 47
 FT DOMAIN 48 65
 FT TRANSMEM 66 92
 FT CYTOPLASMIC (POTENTIAL).

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: AEO04883; AAC08080.1; -
 DR InterPro: IPR000506; ACh_Isomrctse.
 DR Pfam: PF01450; IlyC; 1.
 DR TRIGRAMS: TRIGR00465; IlyC; 1.
 KM Oxidoreductase: Branched-chain amino acid biosynthesis; NADP;
 KM Complete proteome.
 FT ACT_SITE 107 POTENTIAL.
 SQ SEQUENCE 338 AA; 36424 MW; 93CF5F7DA9F61B0A CRC64;

Query Match 33.3%; Score 5; DB 1; Length 338;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRAA 15
 |||||
 Db 269 ESRAA 273

RESULT 64
 LPXD_PSEAE STANDARD; PRT; 353 AA.

AC 09HX6;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE UDP-3-O-(3-hydroxymyristoyl) glucosamine N-acetyltransferase
 DE (EC 2.3.1.1-).
 GN LPXD OR PA3646.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
 RA Hickey M.J., Brinkman F.S.L., Huftagel W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brody L.L., Collier S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Salier M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT opportunistic pathogen.";
 RL Nature 406:959-964 (2000).

CC -1- CATALYTIC ACTIVITY: UDP-3-O-(3-hydroxytetradecanoyl)glucosamine +
 CC (R)-3-hydroxytetradecanoyl-[acyl-carrier protein] = UDP-2,3-bis(3-
 CC hydroxytetradecanoyl)glucosamine + [acyl-carrier protein].
 CC -1- PATHWAY: Lipid A biosynthesis; third step.
 CC -1- SIMILARITY: BELONGS TO THE TRANSFERASE HEXAPEPTIDE REPEAT FAMILY.
 CC LPXD SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AEO04784; AAG07034.1; -
 DR InterPro: IPR001451; Hexapep_transf.
 DR Pfam: PF00132; hexapep; 8.
 DR PROSITE: PS00101; HEXAPEP_TRANSFERASES; 1.
 KM Transferase: Acyltransferase; Lipid A biosynthesis; Lipid synthesis;
 KM Repeat; Complete proteome.
 SQ SEQUENCE 353 AA; 36191 MW; 4B6C4D329F63EA35 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 353;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRA 14
 |||||
 Db 55 PESRA 59

RESULT 65
 GBAF_CAEEL STANDARD; PRT; 356 AA.

AC P91907; Q9BIG1;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Guanine nucleotide-binding protein alpha-15 subunit.
 DE GPA-15 OR M04C7.1.
 GN Caenorhabditis elegans.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;
 RA Cuppen E., Jansen G., Plasterk R.H.A.;
 RT "Interaction analysis of the complete G-alpha subfamily of
 RT heterotrimeric G proteins from Caenorhabditis elegans.";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Kershaw J.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
 CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
 CC SIGNALING SYSTEMS.
 CC -1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS: ALPHA, BETA AND
 CC GAMMA. THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING
 CC SITE.

CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: AT008137; AAC32090.1; -
 DR EMBL: Z83117; CAB0570.2; -
 DR Wormpep: M04C7.1; CE12394.
 DR HSSP: P04896; IAZT.
 DR InterPro: IPR001019; Gprotein_alpha.
 DR Pfam: PF00503; G-alpha.1.
 DR PRINTS: PR00318; GPROTEIN_A.
 DR Prodom: PD000281; Gprotein_alpha; 1.
 KM GTP-binding; Transducer; Multigene family.
 SQ SEQUENCE 356 AA; 40862 MW; 70F37C61031FA04E CRC64;

Query Match 33.3%; Score 5; DB 1; Length 356;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STRES 12
 |||||
 Db 6 STRES 10

RESULT 66
 PE22_RAT

```

RESULT 61
Y05G_BPT4
ID Y05G_BPT4 STANDARD: PRT: 336 AA.
AC P39242: Q96217:
DT 01-FEB-1995 (Rel. 31, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical 38.9 kDa protein in nrdc-mobd intergenic region.
Y05G OR NRDC.11 OR TK.11.
OS Bacteriophage T4.
CC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
CC T4-like viruses.
OX NCBI_TaxID=10655;
RN [1]
RP SEQUENCE FROM N.A.
RA Mzhanava N., Marusich E., Djavakhishvili T., Neitzel J., Peterson S.,
RA Awaya M., Eldermiller J., Canada D., Tracy J., Gallbreath K.,
RA Paddison P., Anderson B., Stidham T., Blatter F., Kutter E.M.;
RT "The 10.7 kb 'nonessential' region of bacteriophage T4 between the
RT genes tk and nrdc: twenty new t4 genes, generally conserved among
RT T-even phages."
RL Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Kutter E., Arisaka F., Kunisawa T., Tsugita A., Mosig G.,
RA Mesyanzhinov V., Ruger W., Stidham T., Thomas E.;
RT "Bacteriophage T4 genome analysis."
RL Submitted (JUL-2000) to the EMBL/Genbank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U76612; AAB26971.1; -.
DR EMBL: AF158101; AAD42629.1; -.
KM Hypothetical protein.
SQ SEQUENCE 336 AA: 38896 MW: 10681F2953A6A9D0 CRC64:

Query Match 33.3%; Score 5; DB 1; Length 336;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 SNPS 12
DB 17 STPS 21

RESULT 62
PANC_STRCO
ID PANC_STRCO STANDARD: PRT: 337 AA.
AC Q9X844: Q9X8N9:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Pantoate-beta-alanine ligase (EC 6.3.2.1) (Pantothenate synthetase)
DE (Pantoate activating enzyme).
GN PANC OR SC03383 OR SCE126.01C OR SCE94.34C.
OS Streptomyces coelicolor.
CC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
CC Actinomycetales; Streptomycetaceae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-A3(2) / M145;
RC MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,

```

```

RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabdinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Bartell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
CC -1- CATALYTIC ACTIVITY: ATP + (R)-pantoate + beta-alanine -> AMP +
CC diphosphate + (R)-pantothenate.
CC -1- PATHWAY: Pantothenate biosynthesis; last step.
CC -1- SIMILARITY: BELONGS TO THE PANTOTHENATE SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL049630; CAB40923.1; -.
DR EMBL: AL049628; CAB40883.1; -.
DR InterPro: IPR003721; Pantoate_ligase.
DR Pfam: PF02569; Pantoate_ligase; 1.
DR TIGRfam: TIGR00018; panc; 1.
KM Pantothenate biosynthesis; Ligase; Complete proteome.
SQ SEQUENCE 337 AA: 35456 MW: 3C948F7014A213C CRC64:

```

```

Query Match 33.3%; Score 5; DB 1; Length 337;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRA 15
DB 246 ESRA 250

RESULT 63
ILVC_PSEAE
ID ILVC_PSEAE STANDARD: PRT: 338 AA.
AC Q9HVA2:
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ketol-acid reductoisomerase (EC 1.1.1.86) (Acetohydroxy-acid
DE isomerase) (alpha-keto-beta-hydroxyacid reductoisomerase).
GN ILVC OR PA4694.
OS Pseudomonas aeruginosa.
CC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
CC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-ATCC 15692 / PA01;
RC MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Ervin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufrange W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Labdig K., Lim R.M.,
RA Smith K.A., Sailer M.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reitzer J., Satter M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RL Nature 406:959-964(2000).
CC -1- CATALYTIC ACTIVITY: (R)-2,3-dihydroxy-3-methylbutanoate + NADP(+)
CC = (S)-2-hydroxy-2-methyl-3-oxobutanoate + NADPH.
CC -1- PATHWAY: Valine and isoleucine biosynthesis; second step.
CC -1- SIMILARITY: BELONGS TO THE KETOL-ACID REDUCTOISOMERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

RESULT 59
ASBL_HUMAN STANDARD; PRT; 335 AA.
ID ASBL_HUMAN
AC Q9Y576; Q9Y576; 41, Created
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Ankyrin repeat and SOCS box containing protein 1 (ASB-1).
GN ASB1 OR KIA1146.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=20564172; PubMed=1111040;
RA Kille B.T., Viney E.M., Willson T.A., Brodnicki T.C., Cancilla M.R.,
RA Herlihy A.S., Croker B.A., Baca M., Nicola N.A., Hilton D.J.,
RA Alexander W.S.;
RT Cloning and characterization of the genes encoding the ankyrin repeat
RL and SOCS box-containing proteins Asb-1, Asb-2, Asb-3 and Asb-4.";
RL Gene 258:31-41(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Kille B.T., Viney E.M., Willson T.A., Brodnicki T.C., Cancilla M.R.,
RA Herlihy A.S., Croker B.A., Baca M., Nicola N.A., Hilton D.J.,
RA Alexander W.S.;
RT Cloning and characterization of the genes encoding the ankyrin repeat
RL and SOCS box-containing proteins Asb-1, Asb-2, Asb-3 and Asb-4.";
RL Gene 258:31-41(2000).
RN [3]
RP SEQUENCE OF 65-335 FROM N.A.
RC TISSUE=Brain;
RA MEDLINE=20039618; PubMed=10574461;
RA Hirosewa M., Nagase T., Ishikawa K.-I., Kikuno R., Nomura N.,
RA Ohara O.;
RT Characterization of cDNA clones selected by the Genemark analysis
RT from size-fractionated cDNA libraries from human brain.";
RL DNA Res. 6:329-336(1999).
CC -1- SIMILARITY: CONTAINS 6 ANK REPEATS.
CC -1- SIMILARITY: CONTAINS 1 SOCS BOX DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF156777; AAD41894.1; -;
DR EMBL: BC014528; AA14528.1; -;
DR EMBL: AB032972; BA86460.1; -;
DR Genew: HGNC:16011; ASB1.
DR MIM: 605758; -;
DR HSSP: Q00420; 1AWC.
DR Interpro: IPR002110; ANK.
DR Interpro: IPR001496; SOCS.
DR Pfam: PF00023; ank; 6.
DR SMART: SM00248; ANK; 2.
DR SMART: SM00253; SOCS; 1.
DR PROSITE: PS50088; ANK_REPEAT; 3.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE: PS50225; SOCS; 1.
KW ANK repeat; Repeat.
FT REPEAT 36 68 ANK 1.
FT REPEAT 77 106 ANK 2.
FT REPEAT 110 139 ANK 3.
FT REPEAT 143 172 ANK 4.
FT REPEAT 191 220 ANK 5.
FT REPEAT 235 265 ANK 6.
FT DOMAIN 286 335 SOCS BOX.
SQ SEQUENCE 335 AA; 37014 MW; 0843C96AD1AF60D9 CRC64;

Query Match

33.3%; Score 5; DB 1; Length 335;

Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 HRSTP 10
Db 143 HRSTP 147

RESULT 60
ASBL_MOUSE STANDARD; PRT; 336 AA.
ID ASBL_MOUSE
AC Q9WV74;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Ankyrin repeat and SOCS box containing protein 1 (ASB-1).
GN ASB1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclerogasthi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. AND TISSUE SPECIFICITY.
RC STRAIN=C57BL/6;
RA Kille B.T., Viney E.M., Willson T.A., Brodnicki T.C., Cancilla M.R.,
RA Herlihy A.S., Croker B.A., Baca M., Nicola N.A., Hilton D.J.,
RA Alexander W.S.;
RT Cloning and characterization of the genes encoding the ankyrin repeat
RL and SOCS box-containing proteins Asb-1, Asb-2, Asb-3 and Asb-4.";
RL Gene 258:31-41(2000).
CC -1- TISSUE SPECIFICITY: Highest expression in testis, spleen and bone marrow.
CC -----
CC -1- SIMILARITY: CONTAINS 6 ANK REPEATS.
CC -1- SIMILARITY: CONTAINS 1 SOCS BOX DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF155352; AAD38808.1; -;
DR HSSP: Q00420; 1AWC.
DR MGD: MG1:1929735; ASB1.
DR Interpro: IPR002110; ANK.
DR Interpro: IPR001496; SOCS.
DR Pfam: PF00023; ank; 6.
DR SMART: SM00248; ANK; 2.
DR SMART: SM00253; SOCS; 1.
DR PROSITE: PS50088; ANK_REPEAT; 3.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE: PS50225; SOCS; 1.
KW ANK repeat; Repeat.
FT REPEAT 37 69 ANK 1.
FT REPEAT 78 107 ANK 2.
FT REPEAT 111 140 ANK 3.
FT REPEAT 144 173 ANK 4.
FT REPEAT 192 221 ANK 5.
FT REPEAT 236 266 ANK 6.
FT DOMAIN 287 336 SOCS BOX.
SQ SEQUENCE 336 AA; 37231 MW; 966119E1B21A5E4F CRC64;

Query Match 33.3%; Score 5; DB 1; Length 336;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 HRSTP 10
Db 144 HRSTP 148


```

RX MEDLINE-91101290; PubMed-1987375;
RA Krappa R., Knebel-Moersdorf D.;
RT "Identification of the very early transcribed baculovirus gene
RL PE-38."
RL J. Virol. 65:805-812(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-C6;
RX MEDLINE-94303173; PubMed-8030224;
RA Ayres M.D., Howard S.C., Kuzio J., Lopez-Ferber M., Possee R.D.;
RT "The complete DNA sequence of Autographa californica nuclear
RL polyhedrosis virus."
RL Virology 202:586-605(1994).
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M62488; AAA46733.1; -
CC EMBL: L22858; AAA66783.1; -
CC PIR: A43681; A43681.
CC InterPro: IPR001841; Znf.Ring.
CC Pfam: PF00097; zf-C3HC4; 1.
CC SMART: SM00184; RING; 1.
CC PROSITE: PS00518; ZF_RING_1; 1.
CC PROSITE: PS50089; ZF_RING_2; 1.
CC Early protein; Zinc-finger; DNA-binding.
CC ZN_FING 86 139
CC FT DOMAIN 228 249 LEUCINE-ZIPPER.
CC CONFLICT 291 291 G -> R (IN REF. 1).
CC SEQUENCE 321 AA; 37425 MW; 4D5ACACB6291A60C CRC64;

Query Match 33.3%; Score 5; DB 1; Length 321;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 HRSTP 10
Db 9 HRSTP 13

RESULT 56
MIAA_RALSO STANDARD; PRT; 323 AA.
ID MIAA_RALSO
AC 08XWB0;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE tRNA delta(2)-isopentenylpyrophosphate transferase (EC 2.5.1.8) (IPP
DE transferase) (isopentenyl-diphosphate:tRNA isopentenyltransferase)
DE (IPPT).
GN MIAA OR RSC2564 OR R500755.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-GM11000;
RX MEDLINE-21681879; PubMed-11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brotlier P., Camus J.C., Catolico L.,
RA Chandler M., Choisne N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavey M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Siguer P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum."
RL Nature 415:497-502(2002).

```

```

CC -1- FUNCTION: Catalyzes the first step in the biosynthesis of 2-
CC methylthio-N6-(delta(2)-isopentenyl)-adenosine (MSI2)I(6A)I)
CC adjacent to the anticodon of several tRNA species (By similarity).
CC -1- CATALYTIC ACTIVITY: Isopentenyl diphosphate + tRNA = diphosphate +
CC tRNA containing 6-isopentenyladenosine.
CC -1- SIMILARITY: BELONGS TO THE IPP TRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: A1646070; CAD16271.1; -
CC InterPro: IPR002627; IPPT.
CC Pfam: PF01715; IPPT; 1.
CC ProDom: PD004674; IPPT; 1.
CC TIGRPFAM: TIGR00174; miaA; 1.
CC Transferase; Nucleotidyltransferase; tRNA processing; ATP-binding;
CC Complete proteome.
CC NE_BIND 16 23
CC SEQUENCE 323 AA; 34966 MW; DF0086716A54E180 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 323;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPE 11
Db 293 RSTPE 297

RESULT 57
HX11_HUMAN STANDARD; PRT; 330 AA.
ID HX11_HUMAN
AC P31314; O75699; Q9HCA0;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Homeobox protein Hox-11 (T-cell leukemia, homeobox 1) (TCL-3 proto-
DE oncogene).
DE TLX1 OR HOX11 OR TCL3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-9200958; PubMed-1681546;
RA Kennedy M.A., Gonzalez-Sarmiento R., Kees U.R., Lampert F.,
RA Dear N., Boehm T., Rabbits T.H.;
RT "HOX11, a homeobox-containing T-cell oncogene on human chromosome
RT 10q24."
RT Proc. Natl. Acad. Sci. U.S.A. 88:8900-8904(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-92007734; PubMed-1717256;
RA Lu M., Gong Z., Shen W., Ho A.D.;
RT "The Tcl-3 proto-oncogene altered by chromosomal translocation in
RT T-cell leukemia codes for a homeobox protein."
RL EMBO J. 10:2905-2910(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE-92319541; PubMed-1352396;
RA Lu M., Zhang N., Ho A.D.;
RT "Genomic organization of the putative human homeobox proto-oncogene
RT HOX-11 (TCL-3) and its endogenous expression in T cells."
RL Oncogene 7:1325-1330(1992).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE-91289163; PubMed-1676542;

```



```

DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Melanoma-associated antigen F1 (MAGE-F1 antigen).
GN MAGEF1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
ON NCI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Stone B.C., Schummer M., Paley P.J., Crawford M., Ford M.,
RA Nelson B.H.,
RT "MAGE-F1, a novel ubiquitously expressed member of the MAGE
RT superfamily identified by SEREX immunoscreening."
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-284 FROM N.A.
RC TISSUE=Kidney;
RA Lucas S., Boon T.;
RT Identification of new genes of the MAGE family."
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
CC -1- TISSUE SPECIFICITY: Ubiquitous.
CC -1- SIMILARITY: CONTAINS 1 MAGE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF295378; AAC30208.1; -
DR EMBL: AF320910; AAC38606.1; -
DR InterPro: IPR002190; MAGE.
DR Pfam: PF01454; MAGE: 1.
DR PROSITE: PSS0838; MAGE: 1.
KM Antigen.
KW DOMAIN.
SQ SEQUENCE 308 AA; 35308 MW; 8FB0092653090580 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 308;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 TPESR 13
| | | | |
Db 4 TPESR 8

RESULT 54
PFLE_ECOLI STANDARD; PRT; 308 AA.
ID PFLE_ECOLI
AC P75794;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative pyruvate formate-lyase 3 activating enzyme (EC 1.97.1.4).
GN Y8Y OR B0824.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
ON NCI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=K12 / MG1655;
RX MEDLINE=9742617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
[2]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map."
RL DNA Res. 3:137-155(1996).
CC -1- FUNCTION: ACTIVATION OF PYRUVATE FORMATE-LYASE 2 UNDER ANAEROBIC
CC CONDITIONS BY GENERATION OF AN ORGANIC FREE RADICAL, USING
CC S-ADENOSYLMETHIONINE AND REDUCED FLAVOPOXIN AS COSUBSTRATES TO
CC PRODUCE 5'-DEOXY-ADENOSINE (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + dihydroflavodoxin +
CC [formate acetyltransferase]-glycine = 5'-deoxyadenosine +
CC methionine + flavodoxin + [formate acetyltransferase]-glycine-2-yl
CC radical.
CC -1- COPACATOR: BINDS 1 4FE-4S CLUSTER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE ORGANIC RADICAL ACTIVATING ENZYMES
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000184; AAC73911.1; -
DR EMBL: D90720; BAA35512.1; -
DR EMBL: D90721; BAA35519.1; -
DR EcoGene: EG13470; yd1y.
DR InterPro: IPR001989; Radical_activat.
DR Pfam: PF02143; Radical_activat. 1.
DR ProDom: PD004758; Radical_activat. 1.
DR PROSITE: PS01087; RADICAL_ACTIVATING. 1.
KM Hypothetical protein; Oxidoreductase; Iron-sulfur; 4Fe-4S;
KW Complete proteome.
FT METAL 34 34 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 38 38 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 41 41 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
SQ SEQUENCE 308 AA; 34139 MW; 21A5C573B57578A23 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 308;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRA 14
| | | | |
Db 44 PESRA 48

RESULT 55
PE38_NPVAC STANDARD; PRT; 321 AA.
ID PE38_NPVAC
AC P23801;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Major immediate early protein (PE-38).
GN PE38.
OS Autographa californica nuclear polyhedrosis virus (ACMNPV).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
ON NCI_TaxID=46015;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=E;

```

[1]
RN SEQUENCE FROM N.A.
RC STRAIN=Fischer; TISSUE=Anterior pituitary;
RX MEDLINE=94043133; PubMed=7901208;
RA Andersen B., Schonemann M.D., Pearse R.V. II, Jenne K., Sugerman J.,
RT Rosenfield M.G.;
RT "Brn-5 is a divergent POU domain factor highly expressed in layer IV
of the neocortex";
RL J. Biol. Chem. 268:23390-23398(1993).
CC -1- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS PREFERENTIALLY TO A
CC VARIANT OF THE OCTAMER MOTIF ('ATGATATAT').
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: IN THE EMBRYO, WIDELY EXPRESSED, WITH HIGHEST
CC LEVELS IN THE DEVELOPING BRAIN AND SPINAL CORD. IN THE ADULT,
CC MOSTLY FOUND IN THE BRAIN, WHERE IT IS DIFFUSELY EXPRESSED WITH
CC THE EXCEPTION OF AN ENRICHMENT IN LAYER IV OF THE NEOCORTEX. ALSO
CC FOUND IN KIDNEY, LUNG, HEART, ADRENAL, SKIN, AND PLACENTA. LOW
CC LEVELS IN SPLEEN, MUSCLE, LIVER, ANTERIOR PITUITARY, TESTIS AND
CC OVARY.
CC -1- DOMAIN: CONTAINS TWO DIRECT REPEATS OF 7 AMINO ACIDS IN THE NH2-
CC TERMINUS.
CC -1- SIMILARITY: BELONGS TO THE POU TRANSCRIPTION FACTOR FAMILY.
CC CLASS=6 SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 1 HOMEBOX DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L23204; -; NOT_ANNOTATED_CDS.
CC HSSP: P14859; 1OCT.
DR InterPro: IPR001356; Homeobox.
DR InterPro: IPR000327; POU_domain.
DR Pfam: PF00046; homeobox; 1.
DR Pfam: PF00157; pou; 1.
DR PRINTS: PR00028; POU_DOMAIN.
DR PRODOM: PD000010; Homeobox; 1.
DR PRODOM: PD000583; POU_domain; 1.
DR SMART: SM00389; HOX; 1.
DR SMART: SM00352; POU; 1.
DR PROSITE: PS00035; POU_1; 1.
DR PROSITE: PS00465; POU_2; 1.
DR PROSITE: PS00027; HOMEBOX_1; FALSE_NEG.
DR PROSITE: PS0071; HOMEBOX_2; 1.
KW Homeobox; DNA-binding; Transcription regulation; Nuclear protein;
KW Repeat.
FT DOMAIN 1 139 GLN/PRO-RICH (BASIC).
FT DOMAIN 11 56 2 X 7 AA REPEATS OF N-A-Q-G-Q-V-I.
FT REPEAT 11 17 1.
FT REPEAT 50 56 2.
FT DOMAIN 140 213 POU.
FT DNA_BIND 234 293 HOMEBOX.
SQ SEQUENCE 301 AA; 32701 MW; 44A50F627EB9F28 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 301;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPES 12
|||||
DB 74 STPES 78

RESULT 52
MTRF_METJA STANDARD; PRT; 303 AA.
AC 058257;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tetrahydromethanopterin S-methyltransferase subunit E (EC 2.1.1.86)
DE (N5-methyltetrahydromethanopterin-coenzyme M methyltransferase
DE subunit E)
GN MTRF OR M0847.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Ullrich T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL Science 273:1058-1073(1996).
CC -1- FUNCTION: THIS ENZYME COMPLEX CATALYZES AN INTERMEDIATE STEP IN
CC METHANOGENESIS, THE FORMATION OF METHYL-COENZYME M AND
CC TETRAHYDROMETHANOPTERIN FROM COENZYME M AND N5-METHYL-
CC TETRAHYDROMETHANOPTERIN. THIS IS A ENERGY-CONSERVING, SODIUM-ION
CC TRANSLLOCATING STEP (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: 5-methyl-5,6,7,8-tetrahydromethanopterin + 2-
CC mercaptoethanesulfonate -> 5,6,7,8-tetrahydromethanopterin + 2-
CC (methylthio)ethanesulfonate.
CC -1- PATHWAY: Methanogenesis.
CC -1- SUBUNIT: COMPOSED OF 8 DIFFERENT SUBUNITS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U67529; AAB98852.1; -
DR TIGR: M0847; -
DR TIGR: M0847; -
KW TIGR: M0847; TIGR01113; mtrE; 1.
KW Transferase; Methyltransferase; Transmembrane; Methanogenesis;
KW Complete proteome.
FT TRANSMEM 3 23 POTENTIAL.
FT TRANSMEM 63 83 POTENTIAL.
FT TRANSMEM 86 106 POTENTIAL.
FT TRANSMEM 133 153 POTENTIAL.
FT TRANSMEM 157 177 POTENTIAL.
FT TRANSMEM 226 246 POTENTIAL.
FT TRANSMEM 266 286 POTENTIAL.
SQ SEQUENCE 303 AA; 31635 MW; 46BA1BD80FECCE2 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 303;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGP 5
|||||
DB 131 SHLGP 135

RESULT 53
MGFL_HUMAN STANDARD; PRT; 308 AA.
ID MGFL_HUMAN
AC 0984V2; Q98425;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)

```

RA Nataka D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
RA Mal'kh A.G., Koonin E.V., Kozlovskiy S.A.;
RT "The complete genome of hyperthermophilic Methanopyrus kandleri AV19
RT and monophyly of archaeal methanogens."
RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
RN [1]
RP SEQUENCE OF 1-89 FROM N.A.
RX MEDLINE=97016827; PubMed=8863453;
RA Noelting J., Elner A., Palmer J.R., Steigewald V.J., Pihl T.D.,
RA Lake J.A., Reeve J.N.;
RT "Phylogeny of Methanopyrus kandleri based on methyl coenzyme M
RT reductase operons."
RL J. Syst. Bacteriol. 46:1170-1173(1996).
CC -1- FUNCTION: THIS ENZYME COMPLEX CATALYZES AN INTERMEDIATE STEP IN
CC METHANOGENESIS, THE FORMATION OF METHYL-COENZYME M AND
CC TETRAHYDROMETHANOPTERIN FROM COENZYME M AND N5-METHYL-
CC TETRAHYDROMETHANOPTERIN. THIS IS A ENERGY-CONSERVING, SODIUM-ION
CC TRANSLLOCATING STEP (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: 5-methyl-5,6,7,8-tetrahydromethanopterin + 2-
CC mercaptoethanesulfonate -> 5,6,7,8-tetrahydromethanopterin + 2-
CC (methylthio)ethanesulfonate.
CC -1- PATHWAY: Methanogenesis.
CC -1- SUBUNIT: COMPOSED OF 8 DIFFERENT SUBUNITS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: Y14428; CAA74774.1; -.
DR EMBL: AE010359; AA001871.1; -.
DR EMBL: U57340; AB02004.1; -.
DR TIGR01113; mtre.1.
KW Transferrase; Methyltransferase; Transmembrane; Methanogenesis;
KW Complete proteome.
FT TRANSMEM 3 23 POTENTIAL.
FT TRANSMEM 59 79 POTENTIAL.
FT TRANSMEM 84 104 POTENTIAL.
FT TRANSMEM 132 152 POTENTIAL.
FT TRANSMEM 160 180 POTENTIAL.
FT TRANSMEM 230 250 POTENTIAL.
FT TRANSMEM 251 271 POTENTIAL.
SQ SEQUENCE 298 AA; 31527 MW; B540A568D1168C4C CRC64;

Query Match * 33.3%; Score 5; DB 1; Length 298;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGP 5
Db 130 SHLGP 134

```

```

RC TISSUE=Skeletal muscle;
RX MEDLINE=94192665; PubMed=7908264;
RA Wey E., Lyons G.E., Schaefer B.W.;
RT "A human POU domain gene, mPOU, is expressed in developing brain and
RT specific adult tissues."
RL Eur. J. Biochem. 220:753-762(1994).
CC -1- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS PREFERENTIALLY TO A
CC VARIANT OF THE OCTAMER MOTIF ('ATGATTAAT') (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: IN THE EMBRYO, EXPRESSED EXCLUSIVELY IN THE
CC DEVELOPING BRAIN, WHEREAS IN THE ADULT ITS EXPRESSION IS
CC RESTRICTED TO BRAIN, HEART, SKELETAL MUSCLE AND LUNG. IN THE
CC BRAIN, THE HIGHEST EXPRESSION LEVELS ARE FOUND IN SPECIFIC CELL
CC LAYERS OF THE CORTEX, THE OLFACTORY BULB, THE HIPPOCAMPUS AND THE
CC CEREBELLUM.
CC -1- SIMILARITY: BELONGS TO THE POU TRANSCRIPTION FACTOR FAMILY.
CC -1- CLASS=6 SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 1 HOMEBOX DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: Z21966; CAA79977.1; -.
DR HSSP: P14859; I0CT.
DR TRANSFAC: T04470; -.
DR Genew: HGNC:9224; POU6F1.
DR InterPro: IPR001356; Homeobox.
DR InterPro: IPR000327; POU domain.
DR Pfam: PF00046; homeobox.1.
DR Pfam: PF00157; pou.1.
DR PRINTS: PR00028; POUDOMAIN.
DR PRODOM: PD000010; Homeobox.1.
DR PRODOM: PD000583; POU domain; 1.
DR SMART: SM00389; Hox.1.
DR SMART: SM00352; POU.1.
DR PROSITE: PS00035; POU.1.1.
DR PROSITE: PS00465; POU.2.1.
DR PROSITE: PS00027; HOMEBOX_1; FALSE_NEG.
DR PROSITE: PS50071; HOMEBOX_2; 1.
KW Homeobox; DNA-binding; Transcription regulation; Nuclear protein.
FT DOMAIN 1 139 GLN/PRO-RICH.
FT DOMAIN 140 213 POU.
FT DNAS_BIND 234 293 HOMEBOX.
SQ SEQUENCE 301 AA; 32645 MW; 9C532BA84ABAB17A CRC64;

Query Match 33.3%; Score 5; DB 1; Length 301;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPES 12
Db 74 STPES 78

```

```

RESULT 50
PO61_HUMAN
ID PO61_HUMAN STANDARD; PRT; 301 AA.
AC Q14863;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE POU domain, class 6, transcription factor 1 (mPOU homeobox protein)
DE (Brain-specific homeobox/POU domain protein 5) (BRN-5 protein).
GN POU6F1 OR mPOU OR BRN5 OR TCFB1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RESULT 51
PO61_RAT
ID PO61_RAT STANDARD; PRT; 301 AA.
AC P56223;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE POU domain, class 6, transcription factor 1 (Brain-specific
DE homeobox/POU domain protein 5) (BRN-5 protein).
GN POU6F1 OR BRN5.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;

```

```

AC P95052;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 50S ribosomal protein L2.
GN RPLB OR RV0704 OR MT0731 OR MRCY210.23.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteriales; Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacterium.
NCBI_TaxID=1773;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sultun J.E., Taylor K., Whitehead S., Barrall B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RA MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sultun J.E., Taylor K., Whitehead S., Barrall B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Esmailava M.D., Salzberg S.L.,
RA Delcher A., Usterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THIS PROTEIN IS A PRIMARY 23S RNA-BINDING PROTEIN. IT
CC HAS PEPTIDYLTRANSFERASE ACTIVITY (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE L2P FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC -----
DR EMBL; 284395; CAB06467.1; -
DR EMBL; AE006966; AAK44962.1; -
DR HSSP; P04257; 1RL2.
DR TIGR; MT0731; -
DR TubercuList; RV0704; -
DR InterPro; IPR002171; Ribosomal_L2.
DR Pfam; PF00181; Ribosomal_L2; 1.
DR TIGRFAMs; TIGR01171; rplB_bact; 1.
DR PROSITE; PS00467; RIBOSOMAL_L2; 1.
KW Ribosomal protein; rRNA-binding; Complete proteome.
SQ SEQUENCE 280 AA; 30577 MW; 456837947E32FFBB CRC64;

```

```

DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Host-inducible protein A.
OS Rhizobium fredii (Sinorhizobium fredii).
OG Plasmid sym.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
NCBI_TaxID=380;
RX MEDLINE=88086864; PubMed=2447061;
RA Sadovsky M.J., Olson E.R., Foster V.E., Kossiak R.M., Verma D.P.S.;
RT "Two host-inducible genes of Rhizobium fredii and characterization of
RT the inducing compound."
RL J. Bacteriol. 170:171-178(1988).
CC -1- FUNCTION: NOT KNOWN.
CC -1- INDUCTION: IN BACTERIAL-PLANT-SYMBIOSIS THIS PROTEIN IS ENCODED
CC BY A BACTERIAL GENE, WHICH IS INDUCIBLE BY PLANT 4',7'-DIHYDROXY-
CC ISOFLAVONE OR DERIVATIVES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC -----
DR EMBL; M19019; AAA26294.1; -
DR PIR; A43663; A43663.
KW Nodulation; Plasmid.
SQ SEQUENCE 295 AA; 33371 MW; 75D034EB7FA38BED CRC64;

```

Query Match 33.3%; Score 5; DB 1; Length 295;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 11 ESRAA 15
Db 166 ESRAA 170

```

RESULT 49
MTRF_METKA STANDARD; PRT; 298 AA.

```

AC 049606;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tetrahydromethanopterin S-methyltransferase subunit E (EC 2.1.1.86)
DE (N5-methyltetrahydromethanopterin-coenzyme M methyltransferase
DE subunit E).
OS Methanopyrus kandleri.
GN MTRF OR MK0656.
OC Methanopyrus kandleri.
OC Archaea; Euryarchaeota; Methanopyrii; Methanopyrales; Methanopyraceae;
OC Methanopyrus.
NCBI_TaxID=2320;
RX MEDLINE=98121200; PubMed=9461302;
RA Harms U., Thauer R.K.;
RT "Identification of the active site histidine in the corrinoid protein
RT MtrA of the energy-conserving methyltransferase complex from
RT Methanobacterium thermoautotrophicum."
RL Eur. J. Biochem. 250:783-788(1997).
RN
RP SEQUENCE FROM N.A.
RC STRAIN=AV19 / DSM 9639;
RC MEDLINE=21927647; PubMed=11930014;
RA Slesarev A.I., Mezhevaya K.V., Makarova K.S., Polushin N.N.,
RA Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,

```

```

RESULT 45
Y0JG_BACSU          STANDARD:      PRT:      275 AA.
ID Y0JG_BACSU
AC P54544.1
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical lipoprotein y0jg precursor.
GN Y0JG.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RX MEDLINE=97124195; PubMed=8969508;
RA Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,
RA Kobayashi Y.;
RT "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of
RT the Bacillus subtilis genome containing the skin element and many
RT sporulation genes."
RL Microbiology 142:3103-3111(1996).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=980044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bortner M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Broillet S., Brusch C.V., Caldwell B., Capuano N.J., Carter N.M.,
RA Choi S.K., Codan J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
RA Denzot F., Devigne K.M., Dusterhoft A., Ehrlich S.D., Emerson P.F.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Gilm S.Y., Glaser P., Goffeau A., Goldthly E.O., Grandi G.,
RA Guspehl G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaetr-blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogawa A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portelle D., Potwilk S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scifone F.,
RA Sekiguchi J., Sekowska A., Setor S.J., Serier P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Yamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Toasato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambuit R., Medier E., Medier H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Zanchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis."
RT Nature 390:249-256(1997).
-1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
(Possible).
-1- SIMILARITY: BELONGS TO THE OXA1 / 60 KDA IMP FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: D84432; BAA12613.1; -
DR EMBL: 299116; CAB14320.1; -
DR Subtilist; Bg11736; y0jg.

```

```

DR InterPro: IPR001708; 60kDa_innemeb.
DR InterPro: IPR000437; Prok_lipoprot.
DR Pfam: PF02096; 60KD_IMP_1.
DR PROSITE: PS00013; PROKAR_LIPOPROTEIN_FALSE_NEG.
KW Hypothetical protein; Membrane; Lipoprotein; Signal;
KM Complete proteome.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 275 HYPOTHETICAL LIPOPROTEIN Y0JG.
FT LIPID 19 19 N-ACYL DIGLYCERIDE (POTENTIAL).
SQ SEQUENCE 275 AA: 30748 MW; 9BAD79A2BA8E53B CRC64;

Query Match
Best Local Similarity 33.3%; Score 5; DB 1; Length 275;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
Db 160 RSTPE 164

RESULT 46
RL2_MYCTU          STANDARD:      PRT:      280 AA.
ID RL2_MYCTU
AC O06047;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 50S ribosomal protein L2.
GN RPLB.
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=BCG;
RX MEDLINE=98065585; PubMed=9402018;
RA Sander P., Pramanathan T., Meier A., Frischkorn K., Bolliger E.C.;
RT "The role of ribosomal RNAs in macroide resistance."
RL Mol. Microbiol. 26:469-480(1997).
-1- FUNCTION: THIS PROTEIN IS A PRIMARY 23S RNA-BINDING PROTEIN. IT
HAS PEPTIDYLTRANSFERASE ACTIVITY (BY SIMILARITY).
-1- SIMILARITY: BELONGS TO THE L2P FAMILY OF RIBOSOMAL PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: Y13328; CAA73675.1; -
DR HSPF: P04257; RL2.
DR InterPro: IPR002171; Ribosomal_L2.
DR Pfam: PF00181; Ribosomal_L2; 1.
DR TIGRfams: TIGR01171; rplB_bact; 1.
DR PROSITE: PS00467; RIBOSOMAL_L2; 1.
KW Ribosomal protein; rRNA-binding.
SQ SEQUENCE 280 AA: 30665 MW; 5F23CBF4C3D4AAAE CRC64;

Query Match
Best Local Similarity 33.3%; Score 5; DB 1; Length 280;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
Db 26 RSTPE 30

RESULT 47
RL2_MYCTU          STANDARD:      PRT:      280 AA.
ID RL2_MYCTU

```



```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RA MEDLINE-98295987; PubMed-9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Elismailer K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Besham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
RA Horsby T., Jagers K., Krogh A., McLean A., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulten J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Uterback T., Weidman J., Khouli H., Gill J., Mikula A.,
RA Bisai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (Apr-2001) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE GNTN FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC
CC EMBL: 277162; CAB00955.1; ALT_INIT.
DR EMBL: AE006952; AAK44737.1; -.
DR TIGR: MT0514; -.
DR Tuberculist: RV0494; -.
DR InterPro: IPR000524; HTH_GntR.
DR Pfam: PF00392; gntR; 1.
DR PRINTS: PR00035; HTHGNTN.
DR SMART: SM00345; HTH_GNTR; 1.
DR PROSITE: PS00043; HTH_GNTR_FAMILY; 1.
DR PROSITE: PS00043; HTH_GNTR_FAMILY; 1.
KW Hypothetical protein; Transcription regulation; DNA-binding;
KW Complete proteome.
FT DNA_BIND 45 64 H-T-H MOTIF (POTENTIAL).
FT SEQUENCE 242 AA; 26014 MW; 3E927CFC4A6FC1C5 CRC64;
SQ
Query Match 33.3%; Score 5; DB 1; Length 242;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 7 RSTPE 11
DB 127 RSTPE 131

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Testis;
RA Ottenwelder B., Obermaier B., Mewes H.-W., Gassenhuber J.,
RA Wiemann S.;
RT Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE-20289799; PubMed-10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,
RA Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordliek G., Hornischer K., Brandt P.,
RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Rieselmann L., Dagand E.,
RA Wehmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA Lehnach H., Reinhardt R., Yaspo M.-L.;
RT "The DNA sequence of human chromosome 21."
RL Nature 403:311-319(2000).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC
CC EMBL: AL117578; CAB56001.1; -.
DR EMBL: AP001754; -. NOT_ANNOTATED_CDS.
DR Hypothetical protein.
KW NON_TER 1
FT SEQUENCE 247 AA; 26427 MW; EE84E0A0AE4BEF4A CRC64;
SQ
Query Match 33.3%; Score 5; DB 1; Length 247;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 7 RSTPE 11
DB 112 RSTPE 116

```

```

RESULT 42
H1L_DROVI STANDARD: PRT; 250 AA.
AC 024704;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Histone H1.1.
GN H1S1.1 OR H1S1 OR 12206.
OS Drosophila virilis (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7244;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Boehm;
RA Nagel S., Grossbach U.;
RT "Histone H1 genes and histone gene clusters in the genus drosophila."
RL J. Mol. Evol. 51:286-298(2000).
RN [2]
RP DISCUSSION OF SEQUENCE.
RX MEDLINE-99132191; PubMed-9933406;
RA Schiemann J.E., Iozovskaya E.R., Strausbaugh L.D.;
RT "Drosophila virilis has atypical kinds and arrangements of histone

```



```

ID COMB_STRCO STANDARD; PRT; 227 AA.
AC 09F3E6;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Probable 2-phosphosulfolactate phosphatase (EC 3.1.3.-).
GN COMB OR SCO/611 OR SC2H2.09.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thompson N.R., James K.D., Harris D.E., Quail M.A., Kleiser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Hang C.-H., Kleiser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).";
RL Nature 417:141-147(2002);
CC -1 CATALYTIC ACTIVITY: 2-phosphosulfolactate = sulfolactate +
phosphate.
CC -1 COFACTOR: Magnesium (By similarity).
CC -1 SIMILARITY: BELONGS TO THE COMB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL450289; CAC16713.1; ALT. INIT.
KW Hydrolyase; Magnesium; Complete proteome.
SQ SEQUENCE 227 AA; 23722 MW; 475C29AF0A51369D CRC64;

Query Match 33.3%; Score 5; DB 1; Length 227;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRA 15
Db 175 ESRA 179

RESULT 39
ID RUM1_SCHPO STANDARD; PRT; 230 AA.
AC P40380; 074373;
DT 01-FEB-1995 (Rel. 31, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Protein rum1 (p25-rum1).
GN RUM1 OR SPBC32F12.09.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=94166876; PubMed=8121488;
RA Moreno S., Nurse P.;
RT "Regulation of progression through the G1 phase of the cell cycle by

```

```

RT the rum1+ gene.";
RL Nature 367:236-242(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgourou J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckart G., Aert R., Robben J., Grymonprez B.,
RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
RA Beer P., Zimmermann W., Medler H., Wandut R., Furnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cervetti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -1 FUNCTION: REGULATOR OF CELL CYCLE G1 PHASE PROGRESSION. PROBABLY
CC INTERACTS WITH CDC2 TO INHIBITS ITS ACTION UNTIL THE CELL MASS FOR
CC START IS REACHED.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X77730; CAA54786.1; -
DR EMBL: AL023796; CAA19370.1; -
DR PIR: S41043; S41043.
KW Cell cycle.
FT CONFLICT 72 73 ML -> IV (IN REF. 1).
SQ SEQUENCE 230 AA; 25288 MW; C5199FE345F7484A CRC64;

Query Match 33.3%; Score 5; DB 1; Length 230;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STRES 12
Db 15 STRES 19

RESULT 40
ID Y494_MYCTU STANDARD; PRT; 242 AA.
AC G11159;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical transcriptional regulator RV0494.
GN RV0494 OR MT0514 OR MYC20G9.20.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Actinomycetales; Corynebacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;

```


CC -1- DOMAIN: THIS PROTEIN HAS FOUR EF-HAND DOMAINS, TWO OF WHICH MAY
CC BE FUNCTIONAL CALCIUM-BINDING SITES.
CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS. BELONGS TO
CC THE EFHS SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: L01583; AAA30171.1; -.
DR PIR: A48567; A48567.
DR HSSP: P02593; ICDM.
DR InterPro: IPR002048; EF-hand.
DR ProDom: PD000012; EF-hand; 1.
DR PROSITE: PS000018; EF_HAND; FALSE_NEG.
KW Calcium-binding; Repeat.
FT CA-BIND 72 84 ANCESTRAL CALCIUM SITE 1 (POTENTIAL).
FT CA-BIND 112 123 EF-HAND 2 (POTENTIAL).
FT CA-BIND 148 159 EF-HAND 3 (POTENTIAL).
FT DOMAIN 184 195 ANCESTRAL CALCIUM SITE 4 (POTENTIAL).
SQ SEQUENCE 208 AA; 23687 MW; 308366D009C127C4 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 208;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 RSTPE 11
Db 170 RSTPE 174

RESULT 36
EFH2-TRICR
ID EFH2-TRICR STANDARD; PRT; 208 AA.
AC PA1049;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE EF-hand protein 5 variant 2 (EFH5) (Calmodulin-ubiquitin associated
DE protein CUB2.8).
OS Trypanosoma cruzi.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5693;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-93149197; PubMed=8381204;
RA Ajloka J., Swindle J.T.;
RT "The calmodulin-ubiquitin associated genes of Trypanosoma cruzi:
RT their identification and transcription";
RL Mol. Biochem. Parasitol. 57:127-136(1993).
CC -1- DOMAIN: THIS PROTEIN HAS FOUR EF-HAND DOMAINS, TWO OF WHICH MAY
CC BE FUNCTIONAL CALCIUM-BINDING SITES.
CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS. BELONGS TO
CC THE EFHS SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: L01584; AAA30172.1; -.
DR HSSP: P02593; ICDM.
DR InterPro: IPR002048; EF-hand.
DR Pfam: PF000036; efhand; 1.
DR ProDom: PD000012; EF-hand; 1.
DR PROSITE: PS000018; EF_HAND; FALSE_NEG.

KW Calcium-binding; Repeat.
FT DOMAIN 72 84 ANCESTRAL CALCIUM SITE 1 (POTENTIAL).
FT CA-BIND 112 123 EF-HAND 2 (POTENTIAL).
FT CA-BIND 148 159 EF-HAND 3 (POTENTIAL).
FT DOMAIN 184 195 ANCESTRAL CALCIUM SITE 4 (POTENTIAL).
SQ SEQUENCE 208 AA; 23671 MW; 15516948A9F567E3 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 208;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 RSTPE 11
Db 170 RSTPE 174

RESULT 37
GTT1-ANOGA
ID GTT1-ANOGA STANDARD; PRT; 209 AA.
AC Q94999;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Glutathione S-transferase 1-1 (EC 2.5.1.18) (GST class-theta).
GN GST1-1.
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
OC Culicidae; Anopheles.
OX NCBI_TaxID=7165;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Suakoko;
RX MEDLINE-97190266; PubMed=9038148;
RA Ranson H., Cornet A.J., Fournier D., Vaughan A., Collins F.H.,
RA Hemingway J.;
RT "Cloning and localization of a glutathione S-transferase class I gene
RT from Anopheles gambiae";
RL J. Biol. Chem. 272:5464-5468(1997).
CC -1- FUNCTION: CONFIGURATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
CC OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
CC -1- CATALYTIC ACTIVITY: RX + glutathione -> HX + R-S-glutathione.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. THETA FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: Z71480; CA96104.1; -.
DR HSSP: P30712; ILJR.
DR InterPro: IPR004046; GST_Cterm.
DR InterPro: IPR004045; GST_Nterm.
DR Pfam: PF00043; GST_C; 1.
DR Pfam: PF02796; GST_N; 1.
DR Transferrase; Multigene family.
SQ SEQUENCE 209 AA; 23529 MW; 874577DC725CC403 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 209;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 ESRAA 15
Db 65 ESRAA 69

RESULT 38
COMB-STRCO

DR PRINTS; PRO0060; RIBOSOMAL16.
 DR TIGR0164; rplp_bact; 1.
 DR PROSITE; PS00586; RIBOSOMAL_L16_1; 1.
 DR PROSITE; PS00701; RIBOSOMAL_L16_2; 1.
 KW Ribosomal protein; rRNA-binding; Complete proteome.
 SQ SEQUENCE 136 AA; 15517 MW; E971587DDCEC50E5 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 136;
 Best Local Similarity 100.0%; Pred. No. 60;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRAA 15
 |||||
 DB 112 ESRAA 116

RESULT 33
 HSPD_BRAJA
 ID HSPD_BRAJA STANDARD; PRT; 151 AA.
 AC 069241;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Small heat shock protein hspd.
 GN HSPD.
 OS Bradyrhizobium japonicum.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Bradyrhizobium group; Bradyrhizobium.
 OX NCBI_TaxId=373;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Narberhaus F., Weighofer W., Fischer H.M., Hennecke H.;
 RT "Identification of the Bradyrhizobium japonicum deep gene as part of
 an operon containing small heat shock protein genes".
 RL Submitted (NOV-1997) to the EMBL/Genbank/DBD databases.
 CC -1- SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20)
 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: AJ003064; CA05835.1; -
 DR InterPro; IPR002068; HSP20.
 DR Pfam; PF00011; HSP20; 1.
 DR PROSITE; PS01031; HSP20; 1.
 KW Heat shock; Multigene family.
 SQ SEQUENCE 151 AA; 17272 MW; FB44EF94FB599EE4 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 151;
 Best Local Similarity 100.0%; Pred. No. 66;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRAA 15
 |||||
 DB 147 ESRAA 151

RESULT 34
 LMBV_CHICK
 ID LMBV_CHICK STANDARD; PRT; 198 AA.
 AC 001636;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Laminin beta-1 chain variant (Laminin beta-1-2 chain) (Fragment).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.
 OX NCBI_TaxId=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Eye;
 RX MEDLINE=93015947; PubMed=1400373;
 RA O'Rear J.J.;
 RT "A novel laminin B1 chain variant in avian eye.";
 RL J. Biol. Chem. 267:20555-20557(1992).
 CC -1- FUNCTION: Binding to cells via a high affinity receptor, laminin
 CC is thought to mediate the attachment, migration, and organization
 CC of cells into tissues during embryonic development by interacting
 CC with other extracellular matrix components.
 CC -1- SUBUNIT: Laminin is a complex glycoprotein, consisting of three
 CC different polypeptide chains (alpha, beta, gamma), which are bound
 CC to each other by disulfide bonds into a cross-shaped molecule
 CC comprising one long and three short arms with globules at each
 CC end.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES (MAJOR
 CC COMPONENT).
 CC -1- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: L00963; AAA9140.1; -
 DR InterPro; IPR001886; LamNT.
 DR InterPro; IPR002049; Laminin_EGF.
 DR Pfam; PF00055; Laminin_Nterm; 1.
 DR ProDom; PD002082; LamNT; 1.
 DR SMART; SM00136; LamNT; 1.
 DR PROSITE; PS01248; LAMININ_TYPE_EGF; PARTIAL.
 KW Extracellular matrix; Cell adhesion; Glycoprotein; Basement membrane;
 KW Laminin EGF-like domain.
 FT NON_TER 1
 FT DOMAIN 74 >198 LAMININ N-TERMINAL (DOMAIN VI).
 FT NON_TER 198
 FT NON_TER 198
 SQ SEQUENCE 198 AA; 21830 MW; 6FD669761892C442 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 198;
 Best Local Similarity 100.0%; Pred. No. 84;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRA 14
 |||||
 DB 34 PESRA 38

RESULT 35
 EPH1_TRYCR
 ID EPH1_TRYCR STANDARD; PRT; 208 AA.
 AC P41048;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE EF-hand protein 5 variant 1 (EPH5) (Calmodulin-ubiquitin associated
 DE protein CUB2.65).
 OS Trypanosoma cruzi.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 OX NCBI_TaxId=5693;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=93149197; PubMed=8381204;
 RA Ajlaka J., Swindie J.T.;
 RT "The calmodulin-ubiquitin associated genes of Trypanosoma cruzi:
 RT their identification and transcription".
 RL Mol. Biochem. Parasitol. 57:127-136(1993).

```

DT 01-NOV-1995 (Rel. 32, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE NADH-ubiquinone oxidoreductase 14.8 kDa subunit (EC 1.6.5.3)
DE (EC 1.6.99.3) (Complex I-14.8kD) (CI-14.8kD).
GN N00-14.8.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN 1)
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-16.
RX STRAIN=74-OR23-1A;
RX MEDLINE=94220045; PubMed=8166654;
RA Azevedo J.E., Ebershorn C., Wener S.;
RT "In organello assembly of respiratory-chain complex I: primary
RT structure of the 14.8 kDa subunit of Neurospora crassa complex I.";
RL Biochem. J. 299:297-302(1994).
CC -1- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
CC CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
CC TO BE UBIQUINONE. THE PRECISE FUNCTION OF THIS SUBUNIT IS UNKNOWN
CC BUT BELIEVED TO BE IMPORTANT.
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -1- CATALYTIC ACTIVITY: NADH + acceptor = NAD(+) + reduced acceptor.
CC -1- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 40 DIFFERENT SUBUNITS.
CC BELONGS TO THE PERIPHERAL ARM OF THE COMPLEX, MOST LIKELY THE
CC GLOBULAR PART.
CC -1- SUBCELLULAR LOCATION: Mitochondrial inner membrane; matrix side.
CC THE IMPORT OF THIS SUBUNIT IS DEPENDENT ON PROTEINACEOUS
CC COMPONENTS TO THE MITOCHONDRIAL OUTER MEMBRANE.
CC -1- SIMILARITY: TO BOVINE COMPLEX I SUBUNIT B14.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X76344; CAA53963.1; -
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
FT INIT_MET 0
FT SEQUENCE 123 AA; 14681 MW; 0367AD9F466EC1AD CRC64;
SQ
Query Match 33.3%; Score 5; DB 1; Length 123;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 RSTPE 11
DB 33 RSTPE 37

```

```

RA Krufit V., Lanka E.;
RT "Nucleotide sequence and organization of genes flanking the transfer
RT origin of promiscuous plasmid RP4.";
RL DNA Seq. 1:303-327(1991).
RN 12)
RP SEQUENCE OF 1-58 FROM N.A.
RX MEDLINE=89184510; PubMed=2538813;
RA Lanka E., Euerste J.P.;
RT "Conjugative transfer of promiscuous IncP plasmids: interaction of
RT plasmid-encoded products with the transfer origin.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:1771-1775(1989).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X54459; CAA38339.1; -
KW EMBL: M25423; AAA26422.1; -
DR PIR: S23003; S23003.
KW Plasmid.
FT INIT_MET 0
FT SEQUENCE 133 AA; 14585 MW; 3D3F3BC1619C2512 CRC64;
SQ
Query Match 33.3%; Score 5; DB 1; Length 133;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 11 ESRRA 15
DB 13 ESRRA 17

```

```

RESULT 32
RIL6_BUCAI STANDARD; PRT; 136 AA.
ID RIL6_BUCAI
AC P57584;
DR 16-OCT-2001 (Rel. 40, Created)
DR 16-OCT-2001 (Rel. 40, Last sequence update)
DR 16-OCT-2001 (Rel. 40, Last annotation update)
DE 50S ribosomal protein L16.
GN RPLP OR B0517.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OS symbiotic bacterium).
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=118099;
RN 1)
RP SEQUENCE FROM N.A.
RX STRAIN=Tokyo 1998;
RX MEDLINE=20445173; PubMed=10993077;
RA Shigenobu S., Matanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids
RT Buchnera sp. APS.";
RL Nature 407:81-86(2000).
CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA AND IS
CC LOCATED AT THE A SITE OF THE PEPTIDYLTRANSFERASE CENTER
CC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE L16P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP001119; BAB13210.1; -
DR InterPro: IPR000114; Ribosomal_L16.
DR Pfam: PF00252; Ribosomal_L16; 1.

```

```

ID VWM_LVX STANDARD; PRT; 108 AA.
AC P27331;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical 11.8 kDa protein (ORF 3).
OS 11ly virus X.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potexvirus.
OX NCBI_TaxID=12194;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90218039; PubMed=2129538;
RA Kemeiink J., van der Vlugt C.I.M., Linthorst H.J.M.,
RA Derks A.F.L.M., Asjes C.J., Bol J.F.;
RT "Homologies between the genomes of a carlavirus (11ly symptomless
RT virus) and a potexvirus (11ly virus X) from 11ly plants.";
RL J. Gen. Virol. 71:917-924(1990).
CC -1- SIMILARITY: TO ORF3 PROTEIN FROM OTHER POTEVIRUSES AND TO 12 kDa
CC PROTEIN FROM CARLAVIRUSES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X15342; CAA33395.1.
DR InterPro: IPR001896; Plant_vir_prot.
DR Pfam: PF01307; Plant_vir_prot. 1.
DR ProDom: PD001561; Plant_vir_prot. 1.
DR Hypothetical protein: Transmembrane.
FT TRANSMEM 72 92 POTENTIAL.
SQ SEQUENCE 108 AA; 11767 MW; B3D69554D57FE5C CRC64;

Query Match 33.3%; Score 5; DB 1; Length 108;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GPHRS 8
Db 61 GPHRS 65

RESULT 28
RBFA_CLOPE STANDARD; PRT; 116 AA.
AC 08XJB9;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ribosome-binding factor A.
GN RBFA OR CPE1685.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=13 / Type A;
RX PubMed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
CC -1- FUNCTION: Associates with free 30S ribosomal subunits (but not
CC with 30S subunits that are part of 70S ribosomes or polysomes).
CC Essential for efficient processing of 16S rRNA. May interact with
CC the 5'terminal helix region of 16S Rna (BY similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).

```

```

CC -1- SIMILARITY: BELONGS TO THE RBFA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP003191; BAB81391.1.
DR InterPro: IPR000238; Rib_bind_facta.
DR Pfam: PF02033; RBFA; 1.
DR ProDom: PD007327; Rib_bind_facta; 1.
DR TIGRPFAM: TIGR00082; rbfA; 1.
DR PROSITE: PS01319; RBFA; 1.
KW rRNA processing; complete proteome.
SQ SEQUENCE 116 AA; 13448 MW; 32DA45EA8902844F CRC64;

Query Match 33.3%; Score 5; DB 1; Length 116;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPE 11
Db 87 RSTPE 91

RESULT 29
Y151_HALHA STANDARD; PRT; 122 AA.
AC P04138;
DT 01-NOV-1986 (Rel. 03, Created)
DT 01-NOV-1986 (Rel. 03, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 13.7 kDa protein in transposable element ISH50.
OS Halobacterium halobium.
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=2242;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83246542; PubMed=6306577;
RA Xu W.-L., Doolittle W.F.;
RT "Structure of the archaeobacterial transposable element ISH50.";
RL Nucleic Acids Res. 11:4195-4199(1983).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X01584; CAB37935.1.
DR PIR: A05114; A05114.
KW Transposable element; Hypothetical protein.
SQ SEQUENCE 122 AA; 13792 MW; C70774DA8969F6E CRC64;

Query Match 33.3%; Score 5; DB 1; Length 122;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 TPESR 13
Db 4 TPESR 8

RESULT 30
NB4M_NEUCR STANDARD; PRT; 123 AA.
ID NB4M_NEUCR
AC P42114;

```

```

RESULT 25
CYC_EUGGR STANDARD; PRT: 102 AA.
ID CYC_EUGGR
AC P00076;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cytochrome c.
OS Euglena gracilis.
OC Eukaryota; Eulenzozoa; Euglenida; Euglenales; Euglena.
OX NCBI_TaxID=3039;
RN [1]
RP SEQUENCE.
RX MEDLINE=76039443; PubMed=170910;
RA Pettigrew G.W., Leaver J.L., Meyer T.E., Ryle A.P.;
RT "Purification, properties and amino acid sequence of atypical
cytochrome c from two protozoa, Euglena gracilis and Crithidia
oncopelti."
RL Blochem. J. 147:291-302(1975).
CC -1- FUNCTION: Electron carrier protein. The oxidized form of the
cytochrome c heme group can accept an electron from the heme group
of the cytochrome c1 subunit of cytochrome oxidase. Cytochrome c
then transfers this electron to the cytochrome oxidase complex.
CC the final protein carrier in the mitochondrial electron-transport
chain.
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -1- PTM: Binds one heme group per molecule.
CC -1- MISCELLANEOUS: COMPARED TO OTHER CYTOCHROME C'S LACKS ONE OF THE
TWO CYSTEINES THAT COVALENTLY BIND THE HEME GROUP.
CC -1- SIMILARITY: Belongs to the cytochrome c family.
DR HSSP: P00004; IREJ.
DR InterPro: IPR000345; CytoC_heme_bind.
DR InterPro: IPR003088; Cyt_C1.
DR Pfam: PF00034; cytochrome_c_1.
DR PRINTS: PR00604; CYTOCHROME_C1.
DR PRODOM: PD000375; Cyt_C1AB_1.
DR PROSITE: PS00190; CYTOCHROME_C; FALSE_NEG.
KW Mitochondrion; Electron transport; Respiratory chain; Heme;
KW Methylation; Acetylation.
FT MOD_RES 1 17 HEME (COVALENT).
FT BINDING 17 18 HEME (AXIAL LIGAND).
FT METAL 18 18 IRON (HEME AXIAL LIGAND).
FT METAL 79 79 IRON (HEME AXIAL LIGAND).
FT MOD_RES 85 85 METHYLATION (TRI-).
SO SEQUENCE 102 AA; 11210 MW; 78414AD1BDA6F84 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 102;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRAA 15
DB 11 ESRAA 15

RESULT 26
YBER_ECOLI STANDARD; PRT: 105 AA.
ID YBER_ECOLI
AC P05848; P77107;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein yber.
GN YBER OR B0637.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE=87030266; PubMed=353535;
RA Asch S., Matsuzawa H., Ishino F., Strominger J.L., Matsushashi M.,
RA Ohta T.;
RT "Nucleotide sequence of the pbpa gene and characteristics of the
RT deduced amino acid sequence of penicillin-binding protein 2 of
RT Escherichia coli K12."
RL Eur. J. Biochem. 160:231-238(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RA Chung E., Allen E., Araujo R., Aparicio A., Davis K., Duncan M.,
RA Federspiel N., Hyman R., Kaiman S., Komp C., Kurd O., Lew H., Lin D.,
RA Namath A., Oefner P., Roberts D., Schramm S., Davis R.W.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horikuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map."
RL DNA Res. 3:137-155(1996).
CC -1- SIMILARITY: STRONG TO H. INFLUENZAE HI0034.
CC -1- SIMILARITY: TO B. SUBTILIS Y0EL.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X04516; CAAC8199.1; ALT_INIT.
DR EMBL: AE000168; AAC73738.1; ALT_INIT.
DR EMBL: U82598; AAB40837.1;
DR EMBL: D90704; BAA35284.1; ALT_INIT.
DR PIR: A24995; QOECPT.
DR EcoGene; EG11255; yber.
DR InterPro: IPR003456; DUF143.
DR InterPro: IPR004394; Iojap.
DR Pfam: PF02410; DUF143.1;
DR TIGRFRAMS: TIGR00090; Iojap; 1.
KW Hypothetical protein; Complete proteome.
SO SEQUENCE 105 AA; 11582 MW; 7C1315607BDB0610 CRC64;

Query Match 33.3%; Score 5; DB 1; Length 105;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRAA 15
DB 58 ESRAA 62

RESULT 27
VMEV_LVX

```

```
FT SITE 398 398 CALCULON ION SELECTIVITY AND PERMEABILITY
FT SITE 994 994 (BY SIMILARITY).
FT SITE 1524 1524 CALCULON ION SELECTIVITY AND PERMEABILITY
FT SITE 1828 1828 (BY SIMILARITY).
FT SITE 212 212 CALCULON ION SELECTIVITY AND PERMEABILITY
FT CARBOHYD 291 212 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1486 1486 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLC 101 120 MISSING (IN ISOFORM 2 AND ISOFORM 3).
FT VARSPLC 1607 1612 MISSING (IN ISOFORM 3).
FT CONFLICT 7 7 A -> S (IN REF. 4).
FT CONFLICT 31 31 E -> K (IN REF. 4).
FT CONFLICT 55 55 A -> S (IN REF. 4).
FT CONFLICT 94 94 L -> V (IN REF. 4).
FT CONFLICT 333 333 M -> V (IN REF. 4).
FT CONFLICT 684 684 A -> V (IN REF. 4).
FT CONFLICT 2080 2080 R -> H (IN REF. 4).
FT CONFLICT 2097 2097 R -> H (IN REF. 3 AND 6).
SQ SEQUENCE 2373 AA; 261205 MW; EBF20A0916450106 CRC64;

Query Match
Best Local Similarity 40.0%; Score 6; DB 1; Length 2373;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 TPESRA 14
DB 2309 TPESRA 2314

RESULT 23
YK03_CABEL STANDARD; PRT; 82 AA.
AC 003561;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein B0464.3 in chromosome III.
GN B0464.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodermidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RC MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kersey J., Kirsten J., Lalister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin A., Roopra A., Saunders D., Shownkeen R.,
RA Sims M., Smailson N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sturton J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterson R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohlschlag P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RL Nature 368:32-38(1994).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
DR PIR; S28280; S28280.
DR Wormpep; B0464.3; CE00017.
KW Hypothetical protein.
SQ SEQUENCE 82 AA; 9427 MW; 10E3CD1E77BC94B8 CRC64;

Query Match
Best Local Similarity 33.3%; Score 5; DB 1; Length 82;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPFR 7
DB 62 LGPFR 66

RESULT 24
PY_DICLA STANDARD; PRT; 97 AA.
AC 09PT98;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Peptide Y precursor.
OS Dicertrichus labrax (European sea bass).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Perciformes; Percoidae;
OC Moronidae; Dicertrichus.
OX NCBI_TaxID=13489;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RC MEDLINE=98292845; PubMed=9629200;
RA Cerdá-Reverter J.M., Martínez Rodríguez G., Zanny S., Carrillo M.,
RA Larhammar D.;
RT "Cloning of neuropeptide Y, peptide YY, and peptide Y from sea bass
RT (Dicertrichus labrax), a marine teleost."
RL Ann. N.Y. Acad. Sci. 839:493-495(1998).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE NPY / PPY / PYY FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC EMBL; AJ005380; CAB64934.1; -.
CC HSSP; P01303; IRON.
DR InterPro: IPR001955; Pancreatc_horm.
DR Pfam: PF00159; hormone3; 1.
DR PRINTS: PD001267; Pancreatc_horm; 1.
DR PRODOM: PD001267; Pancreatc_horm; 1.
DR SMART: SM00309; PAH; 1.
DR PROSITE: PS00265; PANCREATIC_HORMONE_1; 1.
DR PROSITE: PS00276; PANCREATIC_HORMONE_2; 1.
KW Hormone; Cleavage on pair of basic residues; Amidation; Signal;
KW Neuropeptide.
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 64 PEPTIDE Y.
FT PROPEP 68 97 C-TERMINAL EXTENSION.
FT MOD_RES 64 64 AMIDATION (G-65 PROVIDE AMIDE GROUP)
SQ SEQUENCE 97 AA; 10921 MW; EB95943A562DD9C4 CRC64;

Query Match
Best Local Similarity 33.3%; Score 5; DB 1; Length 97;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPE 11
DB 67 RSTPE 71
```


DR Genew: HGNC:8804; PDGFRB.
 DR MIM: 173410; -
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003598; Ig_C2.
 DR InterPro: IPR003600; Ig_Like.
 DR InterPro: IPR001824; RtkinaseIII.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00047; Ig_3.
 DR Pfam: PF00069; pkinase; 1.
 DR ProDom: PD000001; Euk_pkinase; 2.
 DR SMART: SM00410; Ig_Like; 2.
 DR SMART: SM00408; IgC2; 1.
 DR SMART: SM00219; Tyrc; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00240; RECEPTOR_TYR_KIN_III; 1.
 DR Tyrosine-protein kinase: Receptor; Transmembrane; Glycoprotein;
 KW Tyrosine-protein kinase: Receptor; ATP-binding; Immunoglobulin domain;
 KW Signal: Proto-oncogene; Chromosomal translocation.
 FT SIGNAL 1 32
 FT CHAIN 33 1106
 FT DOMAIN 33 531 BETA PLATELET-DERIVED GROWTH FACTOR
 FT TRANSMEM 532 536 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 557 1106 POTENTIAL.
 FT DOMAIN 600 962 CYTOPLASMIC (POTENTIAL).
 FT NP_BIND 606 614 PROTEIN KINASE.
 FT BINDING 634 634 ATP (BY SIMILARITY).
 FT ACT_SITE 826 826 ATP (BY SIMILARITY).
 FT MOD_RES 751 751 PHOSPHORYLATION (AUTO-).
 FT MOD_RES 857 857 PHOSPHORYLATION (AUTO-).
 FT CARBOHYD 45 45 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 89 89 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 215 215 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 354 354 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 371 371 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 468 468 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 479 479 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 241 241 E -> D (IN REF. 2).
 SQ SEQUENCE 1106 AA; 123967 MW; 038C15E531D6E89D CRC64;
 Query Match 40.0%; Score 6; DB 1; Length 1106;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SHLGH 6
 Db 656 SHLGH 661
 RESULT 20
 MLH3_HUMAN STANDARD: PRT; 1453 AA.
 AC 09UHC1; 09UHC0; Q9P292; P49751;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE DNA mismatch repair protein MLH3 (Mult protein homolog 3).
 GN MLH3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RX MEDLINE=20082804; PubMed=10615123;
 RA Lipkin S.M., Wang V., Jacoby R., Banerjee-Basu S., Baxevanis A.D.,

RA Lynch H.T., Elliott R.M., Collins F.S.;
 RT "MLH3: a DNA mismatch repair gene associated with mammalian
 RT microsatellite instability.";
 RL Nat. Genet. 24:27-35(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Fukushima S., Kondo E.;
 RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 1189-1453 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=95319502; PubMed=7596406;
 RA Sherrington R., Rogeev E.I., Liang Y., Rogeev E.A., Levesque G.,
 RA Ikeda M., Chi H., Lin C., Li G., Holman K., Tsuda T., Mar L.,
 RA Foncin J.-F., Bruni A.C., Montes M.P., Sorli S., Rainero I.,
 RA Pines L., Nee L., Chumakov I., Pollen D., Brookes A.,
 RA Sansau P., Polinsky R.J., Masco W., da Silva H.A.R., Haines J.L.,
 RA Pericak-Vance M.A., Tanzi R.E., Roses A.D., Fraser P.E.,
 RA Rommens J.M., St George-Hyslop P.H.;
 RT "Cloning of a gene bearing missense mutations in early-onset familial
 RT Alzheimer's disease.";
 RL Nature 375:754-760(1995).
 RN [4]
 RP VARIANTS HNPCC E-24; S-499; Q-624; C-647; G-817; S-981; S-1007; T-1394
 RP AND K-1451.
 RX MEDLINE=21470348; PubMed=11586295;
 RA Wu Y., Berends M.J.W., Sijmons R.H., Mensink R.G.J., Verlind E.,
 RA Kool K.A., van der Sluis T., Kempinga C., van der Zee A.G.J.,
 RA Hollma H., Buys C.H.C.M., Kleibouker J.H., Hofstra R.M.W.;
 RT "A role for MLH3 in hereditary nonpolyposis colorectal cancer.";
 RL Nat. Genet. 29:137-138(2001).
 CC -1- FUNCTION: PROBABLY INVOLVED IN THE REPAIR OF MISMATCHES IN DNA.
 CC -1- SUBUNIT: HETERODIMER OF MLH1 AND MLH3.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: 1 (SHOWN HERE) AND 2; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: UBQUITOUS.
 CC -1- DISEASE: DEFECTS IN MLH3 MAY BE A CAUSE OF HEREDITARY NON-
 CC POLYPOSIS COLORECTAL CANCER (HNPCC) (LYNCH SYNDROME). HNPCC IS AN
 CC AUTOSOMAL, DOMINANTLY INHERITED DISEASE ASSOCIATED WITH MARKED
 CC INCREASE IN CANCER SUSCEPTIBILITY. IT IS CHARACTERIZED BY A
 CC FAMILIAL PREDISPOSITION TO EARLY ONSET COLORECTAL CARCINOMA (CRC)
 CC AND EXTRA-COLONIC CANCERS OF THE GASTROINTESTINAL, UROLOGICAL AND
 CC FEMALE REPRODUCTIVE TRACTS. HNPCC IS REPORTED TO BE THE MOST
 CC COMMON FORM OF INHERITED COLORECTAL CANCER IN THE WESTERN WORLD.
 CC CANCERS IN HNPCC ORIGINATE WITHIN BENIGN NEOPLASTIC POLYPS TERMED
 CC ADENOMAS. CLINICALLY, HNPCC IS OFTEN DIVIDED INTO TWO SUBGROUPS.
 CC TYPE I: HEREDITARY PREDISPOSITION TO COLORECTAL CANCER, A YOUNG
 CC AGE OF ONSET, AND CARCINOMA OBSERVED IN THE PROXIMAL COLON. TYPE
 CC II: PATIENTS HAVE AN INCREASED RISK FOR CANCERS IN CERTAIN TISSUES
 CC SUCH AS THE UTERUS, OVARY, BREAST, STOMACH, SMALL INTESTINE, SKIN,
 CC AND LARYNX IN ADDITION TO THE COLON. DIAGNOSIS OF CLASSICAL HNPCC
 CC IS BASED ON THE AMSTERDAM CRITERIA: 3 OR MORE RELATIVES AFFECTED
 CC BY COLORECTAL CANCER, ONE A FIRST DEGREE RELATIVE OF THE OTHER
 CC TWO; 2 OR MORE GENERATION AFFECTED; 1 OR MORE COLORECTAL CANCERS
 CC PRESENTING BEFORE 50 YEARS OF AGE; EXCLUSION OF HEREDITARY
 CC POLYPOSIS SYNDROMES. THE TERM "SUSPECTED HNPCC" OR "INCOMPLETE
 CC HNPCC" CAN BE USED TO DESCRIBE FAMILIES WHO DO NOT OR ONLY
 CC PARTIALLY FULFILL THE AMSTERDAM CRITERIA, BUT IN WHOM A GENETIC
 CC BASIS FOR COLON CANCER IS STRONGLY SUSPECTED.
 CC -1- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTL/HEXB FAMILY.
 CC -1- SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 DR EMBL: AF195657; AAF23904.1; -
 DR EMBL: AF195658; AAF23905.1; -
 DR EMBL: AB039667; BAA92353.1; -

```

DR InterPro: IPR003598; Ig_c2.
DR InterPro: IPR003600; Ig_like.
DR InterPro: IPR001824; RTKinasell.
DR InterPro: IPR001245; Tyr_kinase.
DR Pfam: PF00047; ig; 3.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_kinase; 2.
DR SMART: SM00410; Ig_like; 2.
DR SMART: SM00408; Igc2; 1.
DR SMART: SM00219; Tyrc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00240; RECEPTOR_TYR_KIN_III; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR Tyrosine-protein kinase: Receptor; Transmembrane; Glycoprotein;
KW Transferrase: Phosphorylation; ATP-binding; Immunoglobulin domain;
KW Repeat; Signal.
FT SIGNAL 1 31
FT CHAIN 32 1098
FT DOMAIN 32 530 BETA PLATELET-DERIVED GROWTH FACTOR
FT TRANSSEM 531 555 RECEPTOR.
FT DOMAIN 536 1098 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 33 119 CYTOSOLASMIC (POTENTIAL).
FT DOMAIN 128 209 IG-LIKE C2-TYPE DOMAIN 1.
FT DOMAIN 214 310 IG-LIKE C2-TYPE DOMAIN 2.
FT DOMAIN 415 527 IG-LIKE C2-TYPE DOMAIN 3.
FT DOMAIN 599 961 IG-LIKE C2-TYPE DOMAIN 4.
FT BINDING 603 613 PROTEIN KINASE.
FT BINDING 633 633 ATP (BY SIMILARITY).
FT ACT_SITE 825 825 ATP (BY SIMILARITY).
FT DISULFID 53 99 BY SIMILARITY.
FT DISULFID 148 189 POTENTIAL.
FT DISULFID 234 290 POTENTIAL.
FT DISULFID 435 507 POTENTIAL.
FT MOD_RES 750 750 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 856 856 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CARBOHYD 44 44 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 102 102 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 291 291 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 306 306 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 353 353 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 370 370 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 444 444 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 467 467 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 478 478 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1098 AA; 122805 MW; 8D391CAFAC3FC31D CRC64;

Query Match 40.0%; Score 6; DB 1; Length 1098;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 SHLGP 6
DB 655 SHLGP 660

RESULT 19
PGOR_HUMAN
ID PGOR_HUMAN STANDARD; PRT; 1106 AA.
AC P09619;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Beta platelet-derived growth factor precursor (EC 2.7.1.112)
DE (PDGF-R-beta) (CD140b antigen).
DE PDGFRB.
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-88217915; PubMed-2835772;
RA Gronwald R.G.K., Grant F.J., Haldeman B.A., Hart C.E., O'Hara P.J.,
RA Hagen F.S., Ross R., Bowen-Pope D.F., Murray M.J.;
RT "Cloning and expression of a cDNA coding for the human
RT platelet-derived growth factor receptor: evidence for more than one
RT receptor class.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:3435-3439(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-89096941; PubMed-2850496;
RA Claesson-Welsh L., Eriksson A., Momen A., Severinsson L., Ek B.,
RA Oestman A., Besholtz C., Heldin C.-H.;
RT "cDNA cloning and expression of a human platelet-derived growth
RT factor (PDGF) receptor specific for B-chain-containing PDGF
RT molecules.";
RL Mol. Cell. Biol. 8:3476-3486(1988).
RN [3]
RP SEQUENCE OF 548-569 FROM N.A.
RX MEDLINE-97429921; PubMed-928559;
RA Chl K.D., McPhee R.A., Wagner A.S., Dietz J.J., Pantazis P.,
RA Goustin A.S.;
RT "Integration of proviral DNA into the PDGF beta-receptor gene in
RT HTLV-I-infected T-cells results in a novel tyrosine kinase product
RT with transforming activity.";
RL Oncogene 15:1051-1057(1997).
RN [4]
RP SEQUENCE OF 1046-1106 FROM N.A.
RX MEDLINE-89028677; PubMed-2846185;
RA Roberts W.M., Look A.T., Rousset M.F., Sherr C.J.;
RT "Random linkage of human CSF-1 receptor (c-fms) and PDGF receptor
RT genes.";
RL Cell 55:655-661(1988).
RN [5]
RP AUTOPHOSPHORYLATION SITES.
RX MEDLINE-89376563; PubMed-2550144;
RA Kazanietz A., Cooper J.A.;
RT "Autophosphorylation of the PDGF receptor in the kinase insert region
RT regulates interactions with cell proteins.";
RL Cell 58:1121-1133(1989).
RN [6]
RP FUNCTION: THIS RECEPTOR BINDS PLATELET-DERIVED GROWTH FACTOR AND
RP HAS A TYROSINE-PROTEIN KINASE ACTIVITY. THIS RECEPTOR BINDS
RP SPECIFICALLY TO PDGF-B. PHOSPHORYLATES TYR RESIDUES AT THE C-
RP TERMINUS OF PIPN1 CREATING A BINDING SITE FOR THE SH2 DOMAIN OF
RP GRB2.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
CC tyrosine phosphate.
CC -1- SUBUNIT: DIMER OF EITHER ALPHA-ALPHA, BETA-BETA OR ALPHA-BETA
CC SUBUNITS.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- DISEASE: INVOLVED IN A FORM OF CHRONIC MYELOMONOCYTIC LEUKEMIA
CC (CMML) CHARACTERIZED BY ABNORMAL CLONAL MYELOID PROLIFERATION AND
CC BY PROGRESSION TO ACUTE MYELOGENOUS LEUKEMIA (AML). IT IS
CC CHARACTERIZED BY A CHROMOSOMAL TRANSLOCATION T(5;12)(Q33;P13)
CC BETWEEN THE ETS-LIKE PROTEIN TEL AND PDGF RECEPTOR B (PDGFRB).
CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -1- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J03278; AAA60049.1; -
DR EMBL: M21616; AAA36427.1; -
DR EMBL: U33172; AAC51675.1; -
DR PIR: A28206; PF00GB.
DR HSSP: P11362; 1FGK.

```

```

DR EMBL: D13713; BAA02866.1; -
DR PIR: A32402; VGBERB.
DR InterPro: IPR000234; Glycoprot.-B.
DR Pfam: PF00606; Glycoprotein_B; 1.
DR ProDom: PD000693; Glycoprot.-B; 1.
KW Signal; Glycoprotein; Transmembrane.
FT SIGNAL 1 21
FT CHAIN 1 21
FT DOMAIN 22 865
FT TRANSMEM 683 700
FT TRANSMEM 709 729
FT TRANSMEM 732 752
FT DOMAIN 753 865
FT CARBOHYD 184 184
FT CARBOHYD 332 332
FT CARBOHYD 364 364
FT CARBOHYD 406 406
FT CARBOHYD 425 425
FT CARBOHYD 631 631
FT SEQUENCE 865 AA; 98091 MW; B30E93C1AC65CC63 CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 1; Length 865;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRA 14
DB 174 TPESRA 179

RESULT 17
Z03_CANFA
ID Z03_CANFA STANDARD: PRT; 898 AA.
AC 062683;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tight junction protein ZO-3 (Zonula occludens 3 protein) (Zona
DE occludens 3 protein) (Tight junction protein 3).
GN TJP3 OR ZO3.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RA MEDLINE=98198478; PubMed=9531559;
RA Haskins J., Gu L., Wittchen E.S., Hibbard J., Stevenson B.R.;
RT "ZO-3, a novel member of the MAGUK protein family found at the tight
RT junction, interacts with ZO-1 and occludin."
RL J. Cell Biol. 141:199-208(1998).
CC -1- SUBUNIT: INTERACTS WITH OCCLUDIN, CLAUDINS AND ZO-1.
CC -1- SIMILARITY: CONTAINS 3 PDZ/DRH DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 GUANYLATE KINASE-LIKE DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC EMBL: AF023617; AAC39177.1; -.
CC HSSP: P31016; 1BFE.
CC InterPro: IPR000619; Guanylate_kin.
CC InterPro: IPR001478; PDZ.
CC InterPro: IPR001452; SH3.
CC Pfam: PF00595; PDZ; 3.
CC Pfam: PF00625; Guanylate_kin; 1.
CC SMART: SM00072; GUKC; 1.

```

```

DR SMART: SM00228; PDZ; 3.
DR SMART: SM00326; SH3; 1.
DR PROSITE: PS00856; GUANYLATE KINASE 1; FALSE_NEG.
DR PROSITE: PS50052; GUANYLATE_KINASE_2; 1.
DR PROSITE: PS50106; PDZ; 3.
DR PROSITE: PS50002; SH3; 1.
KW Tight junction; SH3 domain; Repeat; Membrane.
FT DOMAIN 11 93
FT DOMAIN 187 264
FT DOMAIN 369 435
FT DOMAIN 464 538
FT DOMAIN 650 750
FT SEQUENCE 898 AA; 98414 MW; 8091D6132DB9F15D CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 1; Length 898;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRAA 15
DB 689 PESRAA 694

RESULT 18
PDDR_MOUSE
ID PDDR_MOUSE STANDARD: PRT; 1098 AA.
AC P05622;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE beta platelet-derived growth factor receptor precursor (EC 2.7.1.112)
DE (PDGF-R-beta).
GN PDGFRB OR PDGFR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Fibroblast;
RA Varden Y., Escobedo J.A., Kuang W.-J., Yang-Feng T.L., Daniel T.O.,
RA Tremble P.M., Chen E.Y., Ando M.E., Haskins R.N., Francke U.,
RA Fried V.A., Ullrich A., Williams L.T.;
RT "Structure of the receptor for platelet-derived growth factor helps
RT define a family of closely related growth factor receptors."
RL Nature 323:226-232(1986).
CC -1- FUNCTION: THIS RECEPTOR BINDS PLATELET-DERIVED GROWTH FACTOR AND
CC HAS A TYROSINE-PROTEIN KINASE ACTIVITY. THIS RECEPTOR BINDS
CC SPECIFICALLY TO PDGF-B.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -1- SUBUNIT: DIMER OF EITHER ALPHA-ALPHA, BETA-BETA OR ALPHA-BETA
CC SUBUNITS.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -1- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC EMBL: X04367; CA27882.1; -.
CC PIR: A25742; PFMSRB.
CC HSSP: P11362; 1FGI.
CC MGD: MGI:97531; Pdgrfb.
CC InterPro: IPR000719; Euk_pkinase.
CC InterPro: IPR003006; IG_MHC.

```

```

CC -1- SIMILARITY: TO OTHER TOMBUSVIRUSES RNA POLYMERASE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M21958; AAR02534.1; ALT-SEQ.
DR PIR: A35315; RRYGCT.
DR InterPro: IPR002584; PV_RdRp.
DR Pfam: PF01615; PV_RdRp; 1.
DR Transferrase: RNA-directed RNA polymerase.
DR CHAIN 1 296 PROTEIN P33.
DR SEQUENCE 817 AA; 92141 MW; ACD829D016762EAB CRC64;

Query Match 40.0%; Score 6; DB 1; Length 817;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRA 14
Db 766 TPESRA 771

RESULT 15
CNBA_MOUSE STANDARD; PRT; 823 AA.
AC 088502;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE High-affinity CAMP-specific and IBMX-insensitive 3',5'-cyclic
DE phosphodiesterase 8A (EC 3.1.4.17) (MMPD8).
GN PDE8A OR PDE8.
OS Mus musculus (Mouse).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Testis;
RX MEDLINE=98338029; Pubmed=9671792;
RA Soderling S.H., Bayuga S.J., Beavo J.A.;
RT Cloning and characterization of a CAMP-specific cyclic nucleotide
RT phosphodiesterase."
RL Proc. Natl. Acad. Sci. U.S.A. 95:8991-8996(1998).
CC -1- FUNCTION: PLAYS A ROLE IN SIGNAL TRANSDUCTION BY REGULATING THE
CC INTRACELLULAR CONCENTRATION OF CYCLIC NUCLEOTIDES. THIS
CC PHOSPHODIESTERASE, WHICH HAS A HIGH AFFINITY FOR CAMP, MAY BE
CC INVOLVED IN MAINTAINING BASAL LEVELS OF THE CYCLIC NUCLEOTIDE
CC AND/OR IN THE CAMP REGULATION OF GERM CELL DEVELOPMENT.
CC -1- CATALYTIC ACTIVITY: Adenosine 3',5'-cyclic phosphate + H(2)O =
CC adenosine 5'-phosphate.
CC -1- COFACTOR: REQUIRES DIVALENT CATIONS. MAGNESIUM OR MANGANESE
CC ARE REQUIRED FOR MAXIMUM ACTIVITY IN VITRO (BY SIMILARITY).
CC -1- ENZYME REGULATION: INHIBITED BY DILIPRIDIMOLE. INSENSITIVE TO
CC SELECTIVE PDE INHIBITOR ROLIPRAM AND TO THE NON-SELECTIVE
CC INHIBITOR, IBMX.
CC -1- PATHWAY: Cyclic nucleotide metabolism.
CC -1- TISSUE SPECIFICITY: HIGHEST LEVELS IN TESTIS > EYE > LIVER >
CC SKELETAL MUSCLE > HEART > 7-DAY EMBRIO > KIDNEY > OVARY > BRAIN.
CC IN THE TESTIS, EXPRESSED SPECIFICALLY IN THE SEMINIFEROUS
CC EPITHELIUM IN A SPATIAL AND TEMPORAL MANNER.
CC -1- DEVELOPMENTAL STAGE: LEVELS OF EXPRESSION DECREASE SOMETIME
CC BETWEEN EMBRYO DAY 7 AND DAY 11. IN THE TESTIS, EXPRESSION
CC RESTRICTED TO MIDDLE AND LATE PACHYTENE SPERMATOCYTES.
CC -1- DOMAIN: COMPOSED OF A C-TERMINAL CATALYTIC DOMAIN CONTAINING TWO
CC PUTATIVE DIVALENT METAL SITES AND AN N-TERMINAL REGULATOR DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE PHOSPHODIESTERASE
CC FAMILY.

```

```

CC -1- SIMILARITY: CONTAINS 1 PAS (PER-ARNT-SIM) DIMERIZATION DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF067806; AAC40194.1; -.
DR MCD: MG1:1277116; Pde8a.
DR InterPro: IPR003607; ME_Pplase_Hdc.
DR InterPro: IPR000014; PAS_domain.
DR InterPro: IPR002073; PDase.
DR Pfam: PF00233; PDase; 1.
DR Pfam: PF00989; PAS; 1.
DR PRINTS: PR00387; PD1ESTERASE1.
DR SMART: SM00471; HDC; 1.
DR SMART: SM00091; PAS; 1.
DR TIGRFAMs: TIGR00229; sensory_box; 1.
DR PROSITE: PS00126; PDASE_1; 1.
DR PROSITE: PS01112; PAS; 1.
DR HYDROLASE; CAMP; Manganese; Magnesium.
FT DOMAIN 216 256 PAS.
FT METAL 526 807 CATALYTIC (BY SIMILARITY).
FT METAL 551 551 MAGNESIUM OR MANGANESE 1 (POTENTIAL).
FT METAL 555 555 MAGNESIUM OR MANGANESE 1 (POTENTIAL).
FT METAL 580 580 MAGNESIUM OR MANGANESE 1 (POTENTIAL).
FT METAL 591 591 MAGNESIUM OR MANGANESE 2 (POTENTIAL).
FT METAL 595 595 MAGNESIUM OR MANGANESE 2 (POTENTIAL).
FT METAL 621 621 MAGNESIUM OR MANGANESE 2 (POTENTIAL).
DR SEQUENCE 823 AA; 93171 MW; 7ED9BEA8EB9BCF2 CRC64;

Query Match 40.0%; Score 6; DB 1; Length 823;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRA 14
Db 704 TPESRA 709

RESULT 16
VGBL_HSYM VGBL_HSYM STANDARD; PRT; 865 AA.
AC P18538;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
GN GB.
OS Marek's disease herpesvirus (strain RB-1B) (MDHV).
OS viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OC NCBI_TaxID=33707;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89293086; Pubmed=2544666;
RA Ross L.J.N., Sanderson M., Scott S.D., Bluns M.M., Doel T., Milne B.;
RA "Nucleotide sequence and characterization of the Marek's disease
RA virus homologue of glycoprotein B of herpes simplex virus."
RL J. Gen. Virol. 70:1789-1804(1989).
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

OX NCBI_TaxID=2097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 33530 / G-37;
 RX MEDLINE=96026346; PubMed=7569993;
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
 RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
 RA Fleischmann J.L., Weidman J.F., Small K.V., Sandusky M., Fierman J.L.,
 RA Nguyen D.T., Uterberty T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
 RA Tomb J.-F., Dougherty B.A., Bort K.F., Hu P.-C., Lueder T.S.,
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
 RT "The minimal gene complement of Mycoplasma genitalium.";
 RL Science 270:397-403(1995).
 RN [2]
 RP SEQUENCE OF 494-702 AND 757-806 FROM N.A.
 RC STRAIN-ATCC 33530 / G-37;
 RX MEDLINE=94075230; PubMed=8253680;
 RA Peterson S.N., Hu P.-C., Bort K.F., Hutchison C.A. III;
 RT "A survey of the Mycoplasma genitalium genome by using random
 RT sequencing.";
 RL J. Bacteriol. 175:7918-7930(1993).
 CC -1- FUNCTION: INVOLVED IN PROTEIN EXPORT. INTERACTS WITH THE SECY/SECE
 CC SUBUNIT. SECA HAS A CENTRAL ROLE IN COUPLING THE HYDROLYSIS OF
 CC ATP TO THE TRANSFER OF PRE-SECRETORY PERIPLASMIC AND OUTER
 CC MEMBRANE PROTEINS ACROSS THE MEMBRANE (BY SIMILARITY).
 CC -1- SUBUNIT: PART OF THE PROKARYOTIC PROTEIN TRANSLOCATION APPARATUS
 CC WHICH COMPRISE SECA, SECB, SECD, SECE, SECF, SECG AND SECY
 CC (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC SIDE OF MEMBRANE
 CC (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SECA FAMILY.
 CC
 CC This swiss-prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U39687; AAC71290.1; -;
 DR EMBL: U01732; AAD10541.1; -;
 DR EMBL: U01743; AAD10553.1; -;
 DR TIGR: MG072; -;
 DR InterPro: IPR001650; Helicase_C.
 DR InterPro: IPR000185; SecA.
 DR Pfam: PF00271; helicase_C; 1.
 DR Pfam: PF01043; SecA_protein; 1.
 DR PRINTS: PRO0906; SECA.
 DR TIGRFAMs: TIGR00963; secA; 1.
 DR PROSITE: PS01312; SECA; 1.
 KW Protein transport; ATP-binding; Membrane; Translocation; Transport;
 KW Complete proteome.
 FT NP_BIND 102 109 ATP (POTENTIAL).
 FT CONFLICT 494 497 IRLG -> YFPS (IN REF. 2).
 FT CONFLICT 560 560 S -> T (IN REF. 2).
 FT CONFLICT 610 618 KORDFLA -> TREVFIS (IN REF. 2).
 FT CONFLICT 624 625 MI -> WS (IN REF. 2).
 FT CONFLICT 629 629 L -> H (IN REF. 2).
 FT SEQUENCE 806 AA; 91584 MW; FDD495AF98B1320F CRC64;
 Query Match 40.0%; Score 6; DB 1; Length 806;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 10 PESRA 15
 DB 255 PESRA 260
 RESULT 13
 RPO_CRV STANDARD; PRT; 817 AA.

AC P17459;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Probable RNA-directed RNA polymerase (EC 2.7.7.48) [Contains: Protein
 DE P33].
 OS Cymbidium ringspot virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tombusviridae;
 CC Tombusvirus.
 OX NCBI_TaxID=12144;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=8936663; PubMed=2771646;
 RA Grieco F., Burgan J., Russo M.;
 RT "The nucleotide sequence of Cymbidium ringspot virus RNA.";
 RL Nucleic Acids Res. 17:6383-6383(1989).
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC {RNA}(N).
 CC -1- MISCELLANEOUS: READTHROUGH OF A TERMINATOR MAY OCCUR: BETWEEN
 CC CODONS FOR 296-LYS AND 297-GLY.
 CC -1- MISCELLANEOUS: READTHROUGH OF THE TERMINATOR CODON UGA OCCURS
 CC BETWEEN THE CODONS FOR 296-LYS AND 297-GLY.
 CC -1- SIMILARITY: TO OTHER TOMBUSVIRUSES RNA POLYMERASE.
 CC
 CC This swiss-prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X15511; CAA33532.1; ALT_SEQ.
 DR PIR: J50268; RVCCR.
 DR InterPro: IPR002564; PV_RdRp.
 DR Pfam: PF01615; PV_RdRp; 1.
 KW Transferase; RNA-directed RNA polymerase.
 FT CHAIN 1 296 PROTEIN P33.
 FT SEQUENCE 817 AA; 91818 MW; 292B7D0C60E208BC CRC64;
 Query Match 40.0%; Score 6; DB 1; Length 817;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 9 TPESRA 14
 DB 766 TPESRA 771
 RESULT 14
 RPO_PBSVC STANDARD; PRT; 817 AA.
 AC P15962;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Probable RNA-directed RNA polymerase (EC 2.7.7.48) [Contains: Protein
 DE P33].
 OS Tomato bushy stunt virus (strain Cherry) (TBSV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tombusviridae;
 CC Tombusvirus.
 OX NCBI_TaxID=12147;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90281577; PubMed=2353450;
 RA Hearne P.O., Knorr D.A., Hillman B.I., Morris T.J.;
 RT "The complete genome structure and synthesis of infectious RNA from
 RT clones of tomato bushy stunt virus.";
 RL Virology 177:141-151(1990).
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC {RNA}(N).
 CC -1- MISCELLANEOUS: READTHROUGH OF THE TERMINATOR CODON UAG OCCURS
 CC BETWEEN THE CODONS FOR 296-LYS AND 297-GLY.

DR InterPro: IPR001108; Presentin.
 DR Pfam: PF01080; Presentin.1.
 DR PRINTS: PR01072; PRESENTIN.
 KW Transmembrane; Glycoprotein.
 FT TRANSMEM 48 68 POTENTIAL.
 FT TRANSMEM 99 119 POTENTIAL.
 FT TRANSMEM 127 147 POTENTIAL.
 FT TRANSMEM 161 181 POTENTIAL.
 FT TRANSMEM 184 204 POTENTIAL.
 FT TRANSMEM 210 230 POTENTIAL.
 FT TRANSMEM 247 267 POTENTIAL.
 FT TRANSMEM 324 394 POTENTIAL.
 FT TRANSMEM 399 419 POTENTIAL.
 FT CARBOHYD 245 245 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 371 371 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 433 AA; 48301 MW; 71CCE3F6BB9C0AF CRC64;

Query Match 40.0%; Score 6; DB 1; Length 493;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 SPESR 13
 |||||
 Db 319 SPESR 324

RESULT 9
 HSF4_HUMAN STANDARD; PRT; 493 AA.
 ID HSF4_HUMAN STANDARD; PRT; 493 AA.
 AC Q9ULV5; Q9ULV6; Q99472;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Heat shock factor protein 4 (HSF 4) (Heat shock transcription factor 4) (HSF 4) (hsf4).
 GN HSF4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RX NCBL_TaxID=9606;
 RP SEQUENCE FROM N.A. (ISOFORM HSF4A).
 RC TISSUE=Heart;
 RA Nakai A., Tanabe M., Kawazoe Y., Inazawa J., Morimoto R.I., Nagata K.;
 RT "HSF4, a new member of the human heat shock factor family which lacks
 RT properties of a transcriptional activator.";
 RL Mol. Cell. Biol. 17:469-481(1997).
 RN [2]
 RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND CHARACTERIZATION.
 RX MEDLINE=99419073; PubMed=10488131;
 RA Tanabe M., Sasaki N., Nagata K., Liu X.-D., Liu P.C.C., Thiele D.J.,
 RA Nakai A.;
 RT "The mammalian HSF4 gene generates both an activator and a repressor
 RT of heat shock genes by alternative splicing.";
 RL J. Biol. Chem. 274:27845-27856(1999).
 CC -1- FUNCTION: DNA-BINDING PROTEIN THAT SPECIFICALLY BINDS HEAT SHOCK
 CC PROMOTER ELEMENTS (HSE). THE HSF4 ISOFORM REPRESENTS TRANSCRIPTION
 CC WHILE THE HSF4B ISOFORM ACTIVATES TRANSCRIPTION.
 CC -1- SUBUNIT: HOMOTRIMER. EXHIBITS CONSTITUTIVE DNA BINDING AND FORMS
 CC TRIMERS EVEN IN THE ABSENCE OF STRESS.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; HSF4A AND HSF4B (SHOWN HERE);
 CC ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN HEART, SKELETAL MUSCLE, AND
 CC BRAIN, AND AT MUCH LOWER LEVELS IN SOME OTHER TISSUES.
 CC -1- SIMILARITY: BELONGS TO THE HSF FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: AB029348; BAA84582.1; -;
 DR EMBL: AB029347; BAA84581.1; -;
 DR EMBL: D87673; BAA13433.1; -;
 DR HSSP: P22813; 1HKT.
 DR TRANSFAC: T04923; -;
 DR Genew: HGNC:5227; HSF4.
 DR MIM: 602438; -;
 DR InterPro: IPR000232; HSF_DNA_bind.
 DR InterPro: IPR002341; HSF_ET5.
 DR Pfam: PF00447; HSF_DNA_bind.1.
 DR PRINTS: PR00056; HSFDOMAIN.
 DR PRODOM: PD001788; HSF_DNA_bind.1.
 DR SMART: SM00415; HSF.1.
 DR PROSITE: PS00434; HSF_DOMAIN.1.
 KW Transcription regulation; Nuclear protein; DNA-binding; Activator;
 KW Repressor; Heat shock; Multigene family; Alternative splicing.
 FT DNA_BIND 18 122
 FT DOMAIN 130 204
 FT DOMAIN 365 390
 FT VARSPLIC 246 320
 FT LPTN/LSPHRARGPISDIPDSPECTRLSPSSDGR
 FT EKGALAIKEPSPGGGAGALAPNEDFCVT -> STY
 FT SLSQRQIMALATGPGAPSSLTSTOKTLHPLRGPGFLPVMW
 FT G (IN ISOFORM HSF4A).
 SQ SEQUENCE 493 AA; 53110 MW; F16389F79EE2BDFE CRC64;

Query Match 40.0%; Score 6; DB 1; Length 493;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 SPESR 13
 |||||
 Db 475 SPESR 480

RESULT 10
 NAB2_HUMAN STANDARD; PRT; 525 AA.
 ID NAB2_HUMAN STANDARD; PRT; 525 AA.
 AC Q15742; Q14797; 076006;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE NGFI-A binding protein 2 (EGR-1 binding protein 2) (Melanoma-
 DE associated delayed early response protein) (MADER protein).
 GN NAB2 OR MADER.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RX NCBL_TaxID=9606;
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE=Placenta;
 RA MEDLINE=96251303; PubMed=8668170;
 RA Svarén J., Severson B.R., Apel E.D., Zimonjic D.B., Popescu N.C.,
 RA Milbrandt J.;
 RT "NAB2, a corepressor of NGFI-A (Egr-1) and Krox20, is induced by
 RT proliferative and differentiative stimuli.";
 RL Mol. Cell. Biol. 16:3545-3553(1996).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RA Gerlinger M., Johnson J.P.;
 RT "Genomic organization of the Mader/NAB2 gene.";
 RL Submitted (May-2000) to the EMBL/Genbank/DBJ databases.
 CC [3]
 CC SEQUENCE OF 50-525 FROM N.A. (ISOFORM 1).
 CC MEDLINE=96243039; PubMed=8649813;
 CC Kirsch K.H., Korradi Y., Johnson J.P.;
 CC "Mader: a novel nuclear protein over expressed in human melanomas.";
 CC Oncogene 12:963-971(1996).
 CC [4]
 CC PARTIAL SEQUENCE FROM N.A. (ISOFORM 3).
 RP

DR TIGR00611; recf; 1.
 DR PROSITE; PS00617; RECF_1; 1.
 DR PROSITE; PS00618; RECF_2; 1.
 KM DNA damage; DNA replication; DNA-binding; SOS response; DNA repair;
 KM ATP-binding; Complete proteome.
 FT N-BIND 30 37 ATP (POTENTIAL).
 SO SEQUENCE 364 AA; 42280 MW; 08FED4025D037373 CRC64;
 Query Match 40.0%; Score 6; DB 1; Length 364;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 LGPHRS 8
 |||||
 DB 254 LGPHRS 259
 RESULT 7
 ID OPS3_DROPS STANDARD; PRT; 382 AA.
 AC P28680;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Opsin Rh3 (inner R7 photoreceptor cells opsin).
 GN RH3.
 OS Drosophila pseudobscura (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7237;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Apple Hill;
 RX MEDLINE=93012921; PubMed=1398053;
 RA Carulli J.P., Harl D.L.;
 RL Genetics 132:193-204(1992).
 CC -1- FUNCTION: VISUAL PIGMENTS ARE THE LIGHT-ABSORBING MOLECULES THAT
 CC MEDIANE VISION. THEY CONSIST OF AN APOPROTEIN, OPSIN, COVALENTLY
 CC LINKED TO CIS-RETINAL.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- PTM: SOME OR ALL OF THE CARBOXYL-TERMINAL SER OR THR RESIDUES MAY
 CC BE PHOSPHORYLATED.
 CC -1- MISCELLANEOUS: EACH DROSOPHILA EYE IS COMPOSED OF 800 FACETS OR
 CC OMMATIDIA. EACH OMMATIDIUM CONTAINS 8 PHOTORECEPTOR CELLS (R1-R8),
 CC THE R1 TO R6 CELLS ARE OUTER CELLS, WHILE R7 AND R8 ARE INNER
 CC CELLS.
 CC -1- MISCELLANEOUS: OPSIN RH3 IS SENSITIVE TO UV LIGHT.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC OPSIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X65879; CAA46710.1; -.
 DR PIR; S24607; S24607.
 DR HSSP; P02699; 1F88.
 DR FlyBase; FBgn0012709; Dpse/Rh3.
 DR InterPro; IPR000276; GPCR_Rhodopsin.
 DR InterPro; IPR001760; Opsin.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PROSITE; PS00237; G_PROTEIN_RECPT_FL_1; 1.
 DR PROSITE; PS00262; G_PROTEIN_RECPT_FL_2; 1.
 DR PROSITE; PS00238; OPSIN; 1.
 KM Photoreceptor; Retinal protein; Transmembrane; Phosphorylation;
 KM Glycoprotein; G-protein coupled receptor; Vision.
 FT DOMAIN 1 56 EXTRACELLULAR.

FT TRANSMEM 57 81 1 (POTENTIAL).
 FT DOMAIN 82 93 CYTOPLASMIC.
 FT TRANSMEM 94 118 2 (POTENTIAL).
 FT DOMAIN 119 132 EXTRACELLULAR.
 FT TRANSMEM 133 152 3 (POTENTIAL).
 FT DOMAIN 153 170 CYTOPLASMIC.
 FT TRANSMEM 171 195 4 (POTENTIAL).
 FT DOMAIN 196 219 EXTRACELLULAR.
 FT TRANSMEM 220 247 5 (POTENTIAL).
 FT DOMAIN 248 283 CYTOPLASMIC.
 FT TRANSMEM 284 307 6 (POTENTIAL).
 FT DOMAIN 308 315 EXTRACELLULAR.
 FT TRANSMEM 316 340 7 (POTENTIAL).
 FT DOMAIN 341 382 CYTOPLASMIC.
 FT CARBOHYD 12 12 N-LINKED (GLCNAC. . .) (PROBABLE).
 FT DISULFID 129 206 POTENTIAL.
 FT BINDING 327 327 RETINAL CHROMOPHORE.
 SO SEQUENCE 382 AA; 42929 MW; AF93922DC0031774 CRC64;
 Query Match 40.0%; Score 6; DB 1; Length 382;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 10 PESRAA 15
 |||||
 DB 362 PESRAA 367
 RESULT 8
 ID PSN1_XENLA STANDARD; PRT; 433 AA.
 AC O12976;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Presentilin alpha.
 GN PS-ALPHA.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Piploidea; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=6335;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=97223465; PubMed=9070286;
 RA Tsujimura A., Yasojima K., Hashimoto-Gotoh T.;
 RT Cloning of Xenopus presentilin alpha and -beta cDNAs and their
 RT differential expression in oogenesis and embryogenesis.";
 RL Biochem. Biophys. Res. Commun. 231:392-396(1997).
 CC -1- FUNCTION: MAY PLAY A ROLE IN NEGATIVE REGULATION OF APOPTOTIC
 CC CASCADES DURING OOGENESIS AND EMBRYOGENESIS, AND IN
 CC DEVELOPMENTALLY MATURED TISSUES SUCH AS BRAIN TISSUE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (by similarity).
 CC -1- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN OVARIES AND TO A LESSER
 CC EXTENT IN TESTIS, INTESTINE, KIDNEY, BRAIN, EYE AND LUNG. WEAK
 CC EXPRESSION IN LIVER AND HEART. PRESENT IN TRACE AMOUNTS IN
 CC SKELETAL MUSCLE.
 CC -1- DEVELOPMENTAL STAGE: ABUNDANT IN EARLY STAGES OF OOGENESIS. THE
 CC EXPRESSION IS RAPIDLY REDUCED BETWEEN MEIOTIC MATURATION AND
 CC FERTILIZATION STAGES.
 CC -1- SIMILARITY: BELONGS TO THE PRESENTILIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D84427; BAA19570.1; -.
 DR MEROPS; A22.001; -.

Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHRST 9
|||||

DB 348 LGPHRST 354

RESULT 5

SYM_MYCPN STANDARD: PRT: 512 AA.
ID SYM_MYCPN STANDARD: PRT: 512 AA.
AC P75091:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Methylonl-tRNA synthetase (EC 6.1.1.10) (Methionine--tRNA ligase)
DE (MetRS).
GN METS OR MPN023 OR MP131.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
RA Hermann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae";
RL Nucleic Acids Res. 24:4420-4449(1996).
RN [2]

RP IDENTIFICATION BY MASS SPECTROMETRY.
RC STRAIN-ATCC 29342 / M129;
RX MEDLINE=21086919; PubMed=11271496;
RA Regula J.T., Ueberle B., Boguth G., Goerg A., Schoelzer M.,
RA Hermann R., Frank R.;
RT "Towards a two-dimensional proteome map of Mycoplasma pneumoniae";
RL Electrophoresis 21:3765-3780(2000).
CC -1- FUNCTION: IT IS PROBABLY ESSENTIAL FOR CELL SURVIVAL, BEING
CC REQUIRED NOT ONLY FOR ELONGATION OF PROTEIN SYNTHESIS BUT ALSO
CC FOR THE INITIATION OF ALL MRNA TRANSLATION THROUGH INITIATOR
CC tRNA(FMET) AMINOACYLATION.

CC -1- CATALYTIC ACTIVITY: ATP + L-methionine + tRNA(Met) = AMP +
CC diphosphate + L-methionyl-tRNA(Met).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC STRONG, TO CYSTEINYL-TRNA SYNTHETASE.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

CC EMBL: AE000015; AAB95779.1; -.
DR HSSP: P23395; IABH.
DR InterPro: IPR002300; tRNA-synt_1a.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002304; tRNA-synt_met.
DR Pfam: PF00133; tRNA-synt_1.1.
DR PRINTS: PR01041; TRNASYNTMET.
DR TIGRfams: TIGR00398; metG; 1.
DR PROSITE: PS00178; AA_TRNA_LIGASE_I; FALSE_NEG.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT SITE 11 21 "HIGH" REGION.
FT SITE 301 305 "KMSK" REGION.
FT BINDING 304 304 ATP (BY SIMILARITY).
SQ SEQUENCE 512 AA; 59263 MW; B5C3861CEC30A6AA CRC64;

Query Match 46.7%; Score 7; DB 1; Length 512;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRAA 15
|||||

DB 180 TPESRAA 186

RESULT 6

RECF_XYLFA STANDARD: PRT: 364 AA.
ID RECF_XYLFA STANDARD: PRT: 364 AA.
AC Q9PHE1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA replication and repair protein recf.
DE RECF OR XF0003.
GN RECF OR XF0003.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OX Xylella.
RN NCBI_TaxID=2371;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN-9a5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carter H.,
RA Coutinho N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Fromme M., Furlan L.R.,
RA Garner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA HO P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitaajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Marques M.V., Miracca E.C., Miyaki C.Y., Montelero-Vilorello C.B.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Montelero-Vilorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A.Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.U.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silva J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
RA Vallada H., Van Sluys M.A., Verjovsky-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Melandis J., Setubal J.C.;

RT "The genome sequence of the plant pathogen Xylella fastidiosa";
RL Nature 406:151-159(2000).
CC -1- FUNCTION: THE RECF PROTEIN IS INVOLVED IN DNA METABOLISM. IT IS
CC REQUIRED FOR DNA REPLICATION AND NORMAL SOS INDUCIBILITY. RECF
CC BINDS PREFERENTIALLY TO SINGLE-STRANDED, LINEAR DNA. IT ALSO SEEMS
CC TO BIND ATP (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE RECF FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

CC EMBL: AE003855; AAF82816.1; -.
DR InterPro: IPR001238; RECF.
DR InterPro: IPR003395; SMC_N.
DR Pfam: PF02463; SMC_N; 1.

RN [2]
 RC SEQUENCE FROM N.A.
 RC STRAIN-Mistar; TISSUE-Brain;
 RX MEDLINE-96255262; PubMed-8710164;
 RA Takahashi H., Murayama M., Takashina A., Mercken M., Nakazato Y.,
 RA Noguchi K., Imahori K.;
 RT "Molecular cloning and expression of the rat homologue of
 RT Presentin-1.";
 RL Neurosci. Lett. 206:113-116(1996).
 CC -1- FUNCTION: MAY PLAY A ROLE IN INTRACELLULAR SIGNALING AND GENE
 CC EXPRESSION OR IN LINKING CHROMATIN TO THE NUCLEAR MEMBRANE.
 CC MAY FUNCTION IN THE CYTOPLASMIC PARTITIONING OF PROTEINS.
 CC -1- SUBUNIT: PREDOMINANTLY HETERODIMER OF A N-TERMINAL (NTF) AND A
 CC C-TERMINAL (CTF) ENDOPROTEOLYTICAL FRAGMENT (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. GOLGI AND
 CC ENDOPLASMIC RETICULUM (BY SIMILARITY).
 CC -1- PTM: PHOSPHORYLATED ON SERINE RESIDUES (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PRESENTIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: D82578; BAA11575.1; -
 DR EMBL: D82363; BAA11564.1; -
 DR MEROPS: A22.001; -
 DR InterPro: IPR001108; Presentin1n.
 DR Pfam: PF01080; Presentin1n; 1.
 DR PRINTS: PRO1072; PRESENTIN1.
 KW Transmembrane; Phosphorylation; Endoplasmic reticulum; Golgi stack.
 FT CHAIN 1 298
 FT DOMAIN 299 468
 FT DOMAIN 1 82
 FT TRANSMEM 83 103
 FT TRANSMEM 104 132
 FT TRANSMEM 133 153
 FT TRANSMEM 154 160
 FT TRANSMEM 161 181
 FT TRANSMEM 182 194
 FT TRANSMEM 195 215
 FT TRANSMEM 216 220
 FT TRANSMEM 221 241
 FT TRANSMEM 242 243
 FT TRANSMEM 244 264
 FT TRANSMEM 265 407
 FT TRANSMEM 408 428
 FT TRANSMEM 433 453
 FT SITE 291 292
 FT SITE 292 293
 FT SITE 293 293
 FT CONFLICT 234 234
 FT CONFLICT 381 381
 FT CONFLICT 468 468
 SQ SEQUENCE 468 AA; 52790 MW; 17CB791E88A1FC0 CRC64;
 Query Match 100.0%; Score 15; DB 1; Length 468;
 Best Local Similarity 100.0%; Pred. No. 1.9e-09;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DT 15-JUL-1999 (rel. 38, last sequence update)
 DT 16-OCT-2001 (rel. 40, last annotation update)
 DE Presentin 1 (PS-1).
 GN PSN1 OR PSN1 OR PS1.
 OS Microcebus murinus (Lesser mouse lemur).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Strepsirhini; Cheirogaleidae;
 OC Microcebus.
 OX NCBI_TaxID=30608;
 RN [1]
 RP SEQUENCE FROM N.A.
 RP TISSUE-Brain;
 RX MEDLINE-97079199; PubMed-8920931;
 RA Calenda A., Mestre-Frances N., Czech C., Pradier L., Bons N.,
 RA Bellis M.;
 RT "Molecular cloning, sequencing, and brain expression of the
 RT presentin 1 gene in Microcebus murinus.";
 RL Biochem. Biophys. Res. Commun. 228:430-439(1996).
 CC -1- FUNCTION: MAY PLAY A ROLE IN INTRACELLULAR SIGNALING AND GENE
 CC EXPRESSION OR IN LINKING CHROMATIN TO THE NUCLEAR MEMBRANE.
 CC MAY FUNCTION IN THE CYTOPLASMIC PARTITIONING OF PROTEINS.
 CC -1- SUBUNIT: PREDOMINANTLY HETERODIMER OF A N-TERMINAL (NTF) AND A
 CC C-TERMINAL (CTF) ENDOPROTEOLYTICAL FRAGMENT (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. GOLGI AND
 CC ENDOPLASMIC RETICULUM (BY SIMILARITY).
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: I-467 (SHOWN HERE) AND I-
 CC 463; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: FOUND PREDOMINANTLY IN NEURONS OF THE
 CC DIFFERENT CORTICAL LAYERS AND HIPPOCAMPUS BUT ALSO IN SUBCORTICAL
 CC STRUCTURES.
 CC -1- PTM: PHOSPHORYLATED ON SERINE RESIDUES (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PRESENTIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: 271333; CA95930.1; -
 DR MEROPS: A22.001; -
 DR InterPro: IPR001108; Presentin1n.
 DR Pfam: PF01080; Presentin1n; 1.
 DR PRINTS: PRO1072; PRESENTIN1.
 KW Transmembrane; Phosphorylation; Endoplasmic reticulum; Golgi stack;
 KW Alternative splicing.
 FT CHAIN 1 298
 FT CHAIN 1 298
 FT DOMAIN 299 467
 FT DOMAIN 1 81
 FT TRANSMEM 82 102
 FT TRANSMEM 103 132
 FT TRANSMEM 133 153
 FT TRANSMEM 154 160
 FT TRANSMEM 161 181
 FT TRANSMEM 182 190
 FT TRANSMEM 191 211
 FT TRANSMEM 212 220
 FT TRANSMEM 221 241
 FT TRANSMEM 242 243
 FT TRANSMEM 244 264
 FT TRANSMEM 265 407
 FT TRANSMEM 408 428
 FT TRANSMEM 433 453
 FT SITE 291 292
 FT SITE 292 293
 FT SITE 293 293
 FT VARSPLIC 26 29
 FT SEQUENCE 467 AA; 52384 MW; D966F2CA7F2975C CRC64;
 SQ SEQUENCE 467 AA; 52384 MW; D966F2CA7F2975C CRC64;
 Query Match 46.7%; Score 7; DB 1; Length 467;

RA Tanabe H., Nishimura T., Li K., St George-Hyslop P.H., Miki T.,
 RA Ogihara T.;
 RT "Three different mutations of presenilin 1 gene in early-onset
 RT Alzheimer's disease families";
 RL Neurosci. Lett. 208:195-198(1996).
 RN [21]
 RP VARIANT AD ASP-135.
 RX MEDLINE-97369208; PubMed-9225696;
 RA Crook R., Ellis R., Shanks M., Thal L.J., Perez-Tur J., Baker M.,
 Query Match 100.0%; Score 15; DB 1; Length 467;
 Best Local Similarity 100.0%; Pred. NO. 1.9e-09;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SHLGPBRSPTESRAA 15
 DB 346 SHLGPBRSPTESRAA 360
 RESULT 2
 PSN1_MOUSE STANDARD; PRT; 467 AA.
 AC P49769; Q9JLP9; (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Presenilin 1 (PS-1) (S182 protein).
 GN PSN1 OR PSNLI OR AD3H.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Brain;
 RX MEDLINE-95319502; PubMed-7596406;
 RA Sherrington R., Rogeev E.I., Liang Y., Rogeev E.A., Levesque G.,
 RA Ikeda M., Chi H., Lin C., Li G., Holman K., Tsuda T., Mar L.,
 RA Foncin J.-F., Bruni A.C., Montesi M.P., Sorbi S., Rahner I.,
 RA Pinessi L., Nee L., Chumakov I., Pollen D., Brookes A.,
 RA Saneau P., Polinsky R.J., Masco W., da Silva H.A.R., Haines J.L.,
 RA Pericak-Vance M.A., Tanzi R.E., Roses A.D., Fraser P.E.,
 RA Rommens J.M., St George-Hyslop P.H.;
 RT "Cloning of a gene bearing missense mutations in early-onset familial
 RT Alzheimer's disease.";
 RL Nature 375:754-760(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-129/SvT;
 RX MEDLINE-97442406; PubMed-9295283;
 RA Mitsuda N., Roses A.D., Vitex M.P.;
 RT "Transcriptional regulation of the mouse presenilin-1 gene";
 RL J. Biol. Chem. 272:23489-23497(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SAM P8; TISSUE-Hippocampus;
 RA Kumar V.B., Vyas K.C., Choudhary V., Franko M., Flood J.F.,
 RA Morley J.E.;
 RT "Molecular cloning and tissue distribution of presenilin-1 in
 RT senescence accelerated mice (SAM P8) mice";
 RL Submitted (May-1999) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: MAY PLAY A ROLE IN INTRACELLULAR SIGNALING AND GENE
 CC EXPRESSION OR IN LINKING CHROMATIN TO THE NUCLEAR MEMBRANE.
 CC -1- SUBUNIT: PREDOMINANTLY HETERODIMER OF A N-TERMINAL (NTF) AND A
 CC C-TERMINAL (CTF) ENDOPROTEOLYTIC FRAGMENT (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: INTERCAL MEMBRANE PROTEIN. GOLGI AND
 CC ENDOPLASMIC RETICULUM (BY SIMILARITY).
 CC -1- PTM: PHOSPHORYLATED ON SERINE RESIDUES (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PRESENILIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L42177; AAC42094.1; -;
 DR EMBL: AF007560; AAB72049.1; -;
 DR EMBL: AF149111; AAF73153.1; -;
 DR MEROPS: A22.001; Psen1.
 DR MGD: MGI:120217; Psen1.
 DR InterPro: IPR001108; Presenilin.
 DR Pfam: PF01080; Presenilin; 1.
 DR PRINTS: PR01072; PRESENILIN.
 KW Transmembrane; Phosphorylation; Endoplasmic reticulum; Golgi stack.
 FT CHAIN 1 298 PRESN1LIN 1 NTF SUBUNIT (BY SIMILARITY).
 FT 299 467 PRESN1LIN 1 CTF SUBUNIT (BY SIMILARITY).
 FT DOMAIN 1 82 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 83 103 POTENTIAL.
 FT DOMAIN 104 132 LUMENAL (POTENTIAL).
 FT TRANSMEM 133 153 POTENTIAL.
 FT DOMAIN 154 160 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 161 181 POTENTIAL.
 FT DOMAIN 182 194 LUMENAL (POTENTIAL).
 FT TRANSMEM 195 215 POTENTIAL.
 FT DOMAIN 221 241 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 242 243 LUMENAL (POTENTIAL).
 FT TRANSMEM 244 264 POTENTIAL.
 FT DOMAIN 265 407 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 408 428 POTENTIAL.
 FT TRANSMEM 433 453 POTENTIAL.
 FT SITE 291 292 ALTERNATIVE PROTEOLYTIC CLEAVAGE SITE (BY
 FT SITE 292 292 ALTERNATIVE PROTEOLYTIC CLEAVAGE SITE (BY
 FT SITE 292 293 ALTERNATIVE PROTEOLYTIC CLEAVAGE SITE (BY
 FT VARIANT 9 9 S -> T (IN STRAIN SAM P8).
 FT VARIANT 40 40 D -> E (IN STRAIN SAM P8).
 FT VARIANT 67 67 E -> CM (IN STRAIN SAM P8).
 FT VARIANT 196 196 V -> L (IN STRAIN SAM P8).
 FT VARIANT 321 322 ER -> RBD (IN STRAIN SAM P8).
 SQ SEQUENCE 467 AA; 52639 MW; D07215B4BD2549 CRC64;
 Query Match 100.0%; Score 15; DB 1; Length 467;
 Best Local Similarity 100.0%; Pred. NO. 1.9e-09;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SHLGPBRSPTESRAA 15
 DB 346 SHLGPBRSPTESRAA 360
 RESULT 3
 PSN1_RAT STANDARD; PRT; 468 AA.
 ID PSN1_RAT
 AC P97887; P97529;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Presenilin 1 (PS-1) (S182 protein).
 GN PSN1 OR PSNLI.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Wistar; TISSUE-Brain;
 RX MEDLINE-97199371; PubMed-9047347;
 RA Taniguchi T., Hashimoto T., Taniguchi R., Shimada K., Kawamata T.,
 RA Yasuda M., Nakai M., Terashima A., Kozumi T., Maeda K., Tanaka C.;
 RT "Cloning of the cDNA encoding rat presenilin-1.";
 RL Gene 186:73-75(1997).

RP SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RX MEDLINE=96160372; PubMed-8574969;
RA Kovacs D.M., Fausett H.J., Page K.J., Kim T.-W., Moir R.D.,
RA Merriam D.E., Hollister R.D., Hallmark O.G., Mancini R.,
RA Palsenstein K.M., Hyman B.T., Tanzl R.E., Wasco W.,
RT "Alzheimer-associated presenilins 1 and 2: neuronal expression in
RT brain and localization to intracellular membranes in mammalian
RT cells.";
RL Nat. Med. 2:224-229(1996).
RN [16]
RP PROCESSING.
RX MEDLINE=97317150; PubMed-9173929;
RA Podlasky M.B., Citron M., Amaratte P., Sherrington R., Xia W.,
RA Zhang J., Diehl T., Levesque G., Fraser P., Haass C., Koo E.H.,
RA Seubert P., St George-Hyslop P., Teplow D.B., Selkoe D.J.,
RT "Presenilin proteins undergo heterogeneous endoproteolysis between
RT Thr291 and Ala299 and occur as stable N- and C-terminal fragments in
RT normal and Alzheimer brain tissue.";
RL Neurobiol. Dis. 3:325-337(1997).
RN [17]
RP FUNCTION, AND MUTAGENESIS OF MET-292.
RX MEDLINE=20014554; PubMed-10545183;
RA Steiner H., Romig H., Pesold B., Philipp U., Baader M., Citron M.,
RA Loetscher H., Jacobsen H., Haass C.,
RT "Amyloidogenic function of the Alzheimer's disease-associated
RT presenilin 1 in the absence of endoproteolysis.";
RL Biochemistry 38:14600-14605(1999).
RN [18]
RP FUNCTION.
RX MEDLINE=20062913; PubMed-10593990;
RA Ray W.J., Yao M., Mumm J., Schroeter E.H., Saffig P., Wolfe M.,
RA Selkoe D.J., Kopan R., Goate A.M.,
RT "Cell surface presenilin-1 participates in the gamma-secretase-like
RT proteolysis of Notch.";
RL J. Biol. Chem. 274:36801-36807(1999).
RN [19]
RP FUNCTION, AND MUTAGENESIS OF ASP-257 AND ASP-385.
RX MEDLINE=99221485; PubMed-10206644;
RA Wolfe M.S., Xia W., Ostaszewski B.L., Diehl T.S., Kimberly W.T.,
RA Selkoe D.J.,
RT "Two transmembrane aspartates in presenilin-1 required for presenilin
RT endoproteolysis and gamma-secretase activity.";
RL Nature 398:513-517(1999).
RN [10]
RP FUNCTION, AND MUTAGENESIS OF ASP-257 AND ASP-385.
RX MEDLINE=20359495; PubMed-10899933;
RA Bercevska O., Jack C., McLean P., Aster J.C., Hicks C., Xia W.,
RA Wolfe M.S., Kimberly W.T., Weinmaster G., Selkoe D.J., Hyman B.T.,
RT "Aspartate mutations in presenilin and gamma-secretase inhibitors both
RT impair notch proteolysis and nuclear translocation with relative
RT preservation of notch1 signaling.";
RL J. Neurochem. 75:583-593(2000).
RN [11]
RP FUNCTION, AND MUTAGENESIS OF LEU-286.
RX MEDLINE=20283925; PubMed-10811883;
RA Kulic L., Walter J., Mulhaup G., Teplow D.B., Baumeister R.,
RA Romig H., Capell A., Steiner H., Haass C.,
RT "Separation of presenilin function in amyloid beta-peptide generation
RT and endoproteolysis of Notch.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:5913-5918(2000).
RN [12]
RP FUNCTION.
RX PubMed-11226248;
RA Bakl L., Marambaud P., Efthimiopoulos S., Georgakopoulos A., Wen P.,
RA Cui W., Shioi J., Koo E., Ozawa M., Friedrich V.L., Robakis N.K.,
RT "Presenilin-1 binds cytoplasmic epithelial cadherin, inhibits
RT cadherin/p120 association, and regulates stability and function of
RT the cadherin/catenin adhesion complex.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:2381-2386(2001).
RN [13]
RP INTERACTION WITH DELTA-2 CATENIN.
RX MEDLINE=9915075; PubMed-10037471;
RA Levesque G., Yu G., Nishimura M., Zhang D.M., Levesque L., Yu H.,
RA Xu D., Liang Y., Rogava E., Ikeda M., Duthie M., Murgolo N., Wang L.,
RA VanderVere P., Bayne M.L., Strader C.D., Rommens J.M., Fraser P.E.,
RT "St George-Hyslop P.;
RT "Presenilins interact with armadillo proteins including
RT neural-specific plakophilin-related protein and beta-catenin.";
RL J. Neurochem. 72:999-1008(1999).
RN [14]
RP REVIEW ON VARIANTS.
RX MEDLINE=97029239; PubMed-8875251;
RA Cruts M., Hendriks L., Van Broeckhoven C.,
RT "The presenilin genes: a new gene family involved in Alzheimer disease
RT pathology.";
RL Hum. Mol. Genet. 5:1449-1455(1996).
RN [15]
RP REVIEW ON VARIANTS.
RX MEDLINE=98180715; PubMed-9521418;
RA Cruts M., van Broeckhoven C.,
RT "Presenilin mutations in Alzheimer's disease.";
RL Hum. Mutat. 11:183-190(1998).
RN [16]
RP VARIANTS AD THR-143 AND ALA-384.
RX MEDLINE=96177673; PubMed-8634711;
RA Cruts M., Backovens H., Wang S.-Y., van Gassen G., Theuns J.,
RA de Jonghe C., Wehnert A., de Voocht J., de Winter G., Gras P.,
RA Bruylant M., Datsun N., Weissenbach J., den Dunnen J.T., Martin J.-J.,
RA Hendriks L., Van Broeckhoven C.,
RT "Molecular genetic analysis of familial early-onset Alzheimer's
RT disease linked to chromosome 14q24.3.";
RL Hum. Mol. Genet. 4:2363-2372(1995).
RN [17]
RP VARIANTS AD L-82; H-115; T-139; R-163; T-231; L-264; V-392 AND Y-410.
RX MEDLINE=96177674; PubMed-8634712;
RA Campion D., Flaman J.-M., Brice A., Hannequin D., Dubois B.,
RA Martin C., Moreau V., Charbonnier F., Didierjean O., Tardieu S.,
RA Penet C., Puel M., Pasquier F., le Dore F., Bellis G., Calenda A.,
RA Heilig R., Martinez M., Mallet J., Bellis M., Clerget-Darpuux F.,
RA Agid Y., Frebourg T.,
RT "Mutations of the presenilin 1 gene in families with early-onset
RT Alzheimer's disease.";
RL Hum. Mol. Genet. 4:2373-2377(1995).
RN [18]
RP VARIANTS AD VAL-260; VAL-285 AND VAL-392.
RX MEDLINE=95379971; PubMed-7651536;
RA Rogava E.I., Sherrington R., Rogava E.A., Levesque G., Ikeda M.,
RA Liang Y., Chi H., Lin C., Holman K., Tsuda T., Mar L., Sorbi S.,
RA Nacmas B., Piacentini S., Amaducci L., Chumakov I., Cohen D.,
RA Lannfelt L., Fraser P.E., Rommens J.M., St George-Hyslop P.H.,
RT "Familial Alzheimer's disease in kindreds with missense mutations in
RT a gene on chromosome 1 related to the Alzheimer's disease type 3
RT gene.";
RL Nature 376:775-778(1995).
RN [19]
RP VARIANTS AD V-139; V-146; Y-163; T-267; A-280 AND G-280.
RX MEDLINE=96024664; PubMed-7550356;
RA Clark R.F., Hutton M., Fuldner R.A., Froelich S., Karan E.,
RA Talbot C., Crook R., London C.L., Pihar G., He C., Korenblatt K.,
RA Martinez A., Wragg M., Busfield F., Behrens M.I., Myers A., Norton J.,
RA Morris J., Mehta N., Pearson C., Lincoln S., Baker M., Duff K.,
RA Zehr C., Perez-Tur J., Houlden H., Ruiz A., Ossa J., Lopez F.,
RA Arcos M., Madridgal L., Collinge J., Humphreys C., Asworth T.,
RA Sarnier S., Fox N.C., Harvey R., Kennedy A., Rogues P.K., Cline R.T.,
RA Phillips C.A., Venter J.C., Forsell L., Axelman K., Lilius L.,
RA Johnston J., Cowburn R., Viltanen M., Windlad B., Kosik K.S.,
RA Hallia M., Poyhonen M., Dickson D., Mann D., Neary D., Snowden J.,
RA Lantos P., Lannfelt L., Rossor M.N., Roberts G.W., Adams M.D.,
RT "The structure of the presenilin 1 (S182) gene and identification of
RT six novel mutations in early onset AD families.";
RL Nat. Genet. 11:219-222(1995).
RN [20]
RP VARIANTS AD PHE-96; ARG-163 AND THR-213.
RX MEDLINE=96310408; PubMed-8733303;
RA Kamino K., Sato S., Sakaki Y., Yoshiwa A., Nishiwaki Y., Takeda H.,

```

107 5 33.3 506 1 SYK_XYLAFA
108 5 33.3 508 1 MORE_BORBU
109 5 33.3 511 1 YAUD_RHISN
110 5 33.3 513 1 YAT3_SCHPO
111 5 33.3 516 1 MERA_XENLA
112 5 33.3 525 1 NAB2_MOUSE
113 5 33.3 543 1 ZN08_HUMAN
114 5 33.3 548 1 MERA_PSEFL
115 5 33.3 552 1 FET4_YEAST
116 5 33.3 560 1 RORG_HUMAN
117 5 33.3 561 1 TARA_MOUSE
118 5 33.3 562 1 HEMA_TJALP
119 5 33.3 562 1 S152_YEAST
120 5 33.3 567 1 GPV_RAT
121 5 33.3 579 1 NH22_CAEEL
122 5 33.3 579 1 SYD_THEMA
123 5 33.3 586 1 KMS_YEAST
124 5 33.3 589 1 STD_MYCLE
125 5 33.3 592 1 FTS2_BARBA
126 5 33.3 593 1 CGP1_CAEEL
127 5 33.3 594 1 SYD_MYCTU
128 5 33.3 596 1 TRE2_RHISP
129 5 33.3 599 1 SYD_SYNY3
130 5 33.3 610 1 YEW6_YEAST
131 5 33.3 611 1 IF4B_HUMAN
132 5 33.3 622 1 VEIN_DROME
133 5 33.3 628 1 V70K_TYMC
134 5 33.3 628 1 V70K_TYMC
135 5 33.3 641 1 FZD1_RAT
136 5 33.3 642 1 FZD1_MOUSE
137 5 33.3 644 1 VP74_NPVOL
138 5 33.3 651 1 PRIA_MYCLE
139 5 33.3 658 1 MUTL_NEIMA
140 5 33.3 658 1 MUTL_NEIMA
141 5 33.3 672 1 ANX6_HUMAN
142 5 33.3 672 1 ANX6_MOUSE
143 5 33.3 672 1 ANX6_RAT
144 5 33.3 680 1 NASP_RABIT
145 5 33.3 681 1 RP3A_MOUSE
146 5 33.3 684 1 RP3A_RAT
147 5 33.3 685 1 NOP4_YEAST
148 5 33.3 701 1 H883_LEITIN
149 5 33.3 708 1 GIT2_MOUSE
150 5 33.3 727 1 NT74_RAT
151 5 33.3 731 1 BGAL_DJACA
152 5 33.3 731 1 BGAL_MALDO
153 5 33.3 733 1 NIBL_HUMAN
154 5 33.3 739 1 VCA1_HUMAN
155 5 33.3 741 1 BSG2_DROME
156 5 33.3 741 1 PA24_BVARE
157 5 33.3 751 1 YK09_YEAST
158 5 33.3 759 1 GIT2_HUMAN
159 5 33.3 765 1 BGLX_ECOLI
160 5 33.3 765 1 BGLX_SALTY
161 5 33.3 775 1 MGDI_MOUSE
162 5 33.3 775 1 MGDI_RAT
163 5 33.3 783 1 MGDI_HUMAN
164 5 33.3 788 1 K6PF_ASPNG
165 5 33.3 787 1 DPOL_HPBDC
166 5 33.3 788 1 DPOL_HPBDB
167 5 33.3 790 1 CAD1_HUMAN
168 5 33.3 792 1 SYFB_CHLPN
169 5 33.3 808 1 SECA_MYCPN
170 5 33.3 815 1 NAHL_RABIT
171 5 33.3 816 1 NAHL_HUMAN
172 5 33.3 832 1 BGAL_ASPOF
173 5 33.3 835 1 BGAL_LYCES
174 5 33.3 845 1 SCPL_MESAU
175 5 33.3 858 1 41_MOUSE
176 5 33.3 859 1 M3KC_HUMAN
177 5 33.3 874 1 ORP5_MOUSE
178 5 33.3 879 1 ORP5_HUMAN
179 5 33.3 880 1 LYTD_BACSU

```

```

09peb6 xylella fas
05119 borrellia bu
P55504 rhizobium s
010150 schizosacch
003414 xenopus lae
061127 mus musculu
P17908 homo sapien
051772 pseudomonas
P40988 saccharomyc
P51449 homo sapien
0994w3 mus musculu
P03451 influenza a
P36429 saccharomyc
008770 rattus norv
009587 caenorhabd1
0941f4 thermotoga
003533 saccharomyc
P36429 mycobacteri
031314 bartonella
018905 caenorhabd1
050649 mycobacteri
053338 rhizobium s
P73851 synechocyst
P40024 saccharomyc
P23588 homo sapien
094918 drosophila
P10357 turnip yell
P28478 turnip yell
008463 rattus norv
070421 mus musculu
010365 oxygia pseu
09ccq3 mycobacteri
091cs2 neisseria m
094jct2 neisseria m
P08133 homo sapien
P14624 mus musculu
P48037 rattus norv
P27123 oryctolagus
P47708 mus musculu
P47709 rattus norv
P37838 saccharomyc
025493 leishmania
0911q2 mus musculu
P31622 rattus norv
Q00662 dianthus ca
P48981 malus domes
096fai1 homo sapien
P19320 homo sapien
P11329 drosophila
P50392 brachydanio
P36124 saccharomyc
Q14161 homo sapien
P33363 escherichia
056078 salmonella
09qyh6 mus musculu
09es73 rattus norv
09y5v3 homo sapien
P78985 aspergillus
P30028 duck hepati
P17192 duck hepati
013634 homo sapien
Q927W0 chlamydia p
P19634 homo sapien
P23791 oryctolagus
P45882 asparagus o
P48880 lycopersico
Q60563 mesocricetu
P48193 mus musculu
012852 homo sapien
09he64 mus musculu
09hox9 homo sapien
P39848 bacillus su

```

```

180 5 33.3 888 1 M3KC_MOUSE
181 5 33.3 888 1 M3KC_RAT
182 5 33.3 928 1 KINH_NEUCR
183 5 33.3 936 1 MSH4_HUMAN
184 5 33.3 940 1 UVR4_ECO57
185 5 33.3 940 1 UVR4_ECOLI
186 5 33.3 943 1 UVR4_PASMU
187 5 33.3 955 1 KINL_LEICH
188 5 33.3 965 1 YNC3_YEAST
189 5 33.3 973 1 TRP5_YEAST
190 5 33.3 974 1 TRP4_MOUSE
191 5 33.3 974 1 TRP5_RABIT
192 5 33.3 975 1 TRP5_MOUSE
193 5 33.3 977 1 TRP4_HUMAN
194 5 33.3 977 1 TRP4_RAT
195 5 33.3 981 1 TRP4_BOVIN
196 5 33.3 1013 1 A60D_DROME
197 5 33.3 1021 1 MANA_RHOMR
198 5 33.3 1026 1 EX5B_CHLTR
199 5 33.3 1036 1 YC35_YEAST
200 5 33.3 1038 1 ITR4_HUMAN

```

ALIGNMENTS

RESULT 1

PSNL_HUMAN STANDARD; PRT; 467 AA.

```

AC P49768; Q14762; Q15719; Q15720;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Presentin 1 (PS-1) (S182 protein).
GN PSEN1 OR PSNL1 OR AD3 OR PST1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid:9606;

```

[1] SEQUENCE FROM N.A., AND VARIANTS AD (ISOFORMS I-467 AND I-463).

```

RC TISSUE-Brain;
RA MEDLINE-95319502; PubMed-7596406;
RA Sherrington R., Rogaeve E.I., Liang Y., Rogaeve E.A., Levesque G.,
RA Ikeda M., Chi H., Lin C., Li G., Holman K., Tsuda T., Mar L.,
RA Foncin J.-F., Bruni A.C., Montesi M.P., Sorbi S., Rainero I.,
RA Pinessi L., Nee L., Chumakov I., Pollen D., Brookes A.,
RA Sansau P., Polinsky R.J., Masco W., da Silva H.A.R., Haines J.L.,
RA Perleak-Vance M.A., Tanzi R.E., Roses A.D., Fraser P.E.,
RA Rommens J.M., St George-Hyslop P.H.;
RA "Cloning of a gene bearing missense mutations in early-onset familial
RA Alzheimer's disease";
RA Nature 375:754-760(1995).

```

[2] SEQUENCE FROM N.A. (ISOFORMS I-463 AND I-374).

```

RC TISSUE-Blood, and Brain;
RA MEDLINE-96193901; PubMed-8641442;
RA Sahara N., Yahagi Y.-I., Takagi H., Kondo T., Okochi M., Usami M.,
RA Shirasawa T., Mori H.;
RA "Identification and characterization of presentin I-467, I-463 and
RA I-374";
RA FEBS Lett. 381:7-11(1996).

```

[3] SEQUENCE FROM N.A.

```

RA Rowen L., Madan A., Qin S., Abbasi N., Dors M., Ratcliffe A.,
RA Madan A., Dickhoff R., Shaffer T., James R., Lasky S., Hood L.;
RT "Complete sequence of the gene for presentin I.";
RT Submitted (NOV-1998) to the EMBL/Genbank/DBJ databases.

```

[4] SEQUENCE OF 1-113 FROM N.A.

```

RA Tsujimura A., Hashimoto-Gotoh T.;
RA Submitted (MAR-1996) to the EMBL/Genbank/DBJ databases.

```

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2003, 13:15:09 ; Search time 21 seconds

(without alignments)

29.626 Million cell updates/sec

Title: US-09-689-159a-2_COPY_346_360

Perfect score: 15

Sequence: 1 SHUGPSTRPESRAA 15

Scoring table: OLIGO

Gapop 60.0 , Gapept 60.0

Searched: 112892 seqs, 41476328 residues

Word size : 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 200 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15	100.0	467	1	PSN1_HUMAN
2	15	100.0	467	1	PSN1_MOUSE
3	15	100.0	467	1	PSN1_RAT
4	7	46.7	467	1	PSN1_MOUSE
5	7	46.7	512	1	SYM_MYCPN
6	6	40.0	364	1	RECE_XYLFA
7	6	40.0	382	1	OPS3_DROPS
8	6	40.0	433	1	PSN1_XENLA
9	6	40.0	433	1	HSF4_HUMAN
10	6	40.0	525	1	NAB2_HUMAN
11	6	40.0	660	1	CIOI_SQUAC
12	6	40.0	806	1	SECA_MYCGE
13	6	40.0	817	1	RRPO_CRY
14	6	40.0	817	1	RRPO_TBSVC
15	6	40.0	823	1	CN8A_MOUSE
16	6	40.0	865	1	VGLB_HSYMD
17	6	40.0	898	1	ZO3_CANPA
18	6	40.0	1098	1	PGDR_MOUSE
19	6	40.0	1106	1	PGDR_HUMAN
20	6	40.0	1453	1	MLH3_HUMAN
21	6	40.0	1711	1	CHDI_MOUSE
22	6	40.0	2373	1	CCAH_HUMAN
23	5	33.3	82	1	YKDS_CAEEL
24	5	33.3	97	1	PY_DICLA
25	5	33.3	102	1	CYC_EUGER
26	5	33.3	105	1	YBEB_ECOLI
27	5	33.3	108	1	VMEH_LVX
28	5	33.3	116	1	RBFA_CLOPE
29	5	33.3	122	1	YTSI_HAIIA
30	5	33.3	123	1	NB4M_NEUCR
31	5	33.3	133	1	TRK4_ECOLI
32	5	33.3	136	1	RL16_BUCAI
33	5	33.3	151	1	HSPD_BRAJA

34	5	33.3	198	1	LMBY_CHICK	001636 gallus gall
35	5	33.3	208	1	EFH2_TRYCR	P41048 trypanosoma
36	5	33.3	208	1	EFH2_TRYCR	P41049 trypanosoma
37	5	33.3	209	1	GTI1_ANOGA	Q94999 anopheles g
38	5	33.3	227	1	COMB_STRCO	Q91366 streptomyces
39	5	33.3	230	1	RDML_SCHPO	P40380 schizosacch
40	5	33.3	242	1	CU94_MYCTU	011159 mycobacteri
41	5	33.3	247	1	CU90_HUMAN	Q9Ufm2 homo sapien
42	5	33.3	250	1	H11_DROVI	Q24704 drosophila
43	5	33.3	253	1	TA4_EIMTE	P13399 elmeria ten
44	5	33.3	274	1	YYVE_CAEEL	Q93834 caenorhabdl
45	5	33.3	275	1	YQUG_BACSU	P54544 bacillus su
46	5	33.3	280	1	RL2_MYCBO	006047 mycobacteri
47	5	33.3	280	1	RL2_MYCCTU	P95052 mycobacteri
48	5	33.3	285	1	H1A_RHIFR	P12779 rhizobium f
49	5	33.3	288	1	MTRF_METKA	Q49606 methanopyru
50	5	33.3	301	1	PO61_HUMAN	Q14863 homo sapien
51	5	33.3	301	1	PO61_RAT	P56223 rattus norv
52	5	33.3	303	1	MTRF_METJA	Q58257 methanococc
53	5	33.3	308	1	MGFI_HUMAN	Q9Hav2 homo sapien
54	5	33.3	308	1	PELE_ECOLI	P23801 escherichia
55	5	33.3	321	1	PE36_NPVAC	P23801 autographa
56	5	33.3	323	1	M1AA_RALSO	Q8XW00 raietonia s
57	5	33.3	330	1	HX11_HUMAN	P31314 homo sapien
58	5	33.3	332	1	HX11_MOUSE	P43345 mus musculu
59	5	33.3	335	1	ASB1_HUMAN	Q9Y576 homo sapien
60	5	33.3	336	1	ASB1_MOUSE	Q9Y574 mus musculu
61	5	33.3	336	1	Y056_BPT4	P39242 bacterioph
62	5	33.3	337	1	PANC_STRCO	Q9X844 streptomyces
63	5	33.3	338	1	ILVC_PSEAE	Q9Hav2 pseudomonas
64	5	33.3	353	1	LEPD_PSEAE	Q9Hxy6 pseudomonas
65	5	33.3	356	1	GBAF_CAEEL	P91907 caenorhabdl
66	5	33.3	357	1	PE22_RAT	Q62928 rattus norv
67	5	33.3	358	1	PE22_HUMAN	P43316 homo sapien
68	5	33.3	363	1	HMR2_DUGTI	Q00401 dugesia clg
69	5	33.3	374	1	RECE_RHIME	P56903 rhizobium m
70	5	33.3	379	1	PANE_YEAST	P38787 saccharomyc
71	5	33.3	379	1	RECE_RHILLO	Q98Bh1 rhizobium l
72	5	33.3	380	1	FOS_MOUSE	P01101 mus musculu
73	5	33.3	381	1	FOS_MSVFB	P01102 fbj murine
74	5	33.3	387	1	MAUG_PARDE	Q51658 paracoccus
75	5	33.3	388	1	VEVY_MCV1	P26579 molluscum c
76	5	33.3	388	1	VEVY_MCV2	P25392 molluscum c
77	5	33.3	402	1	CENC_SHEEP	P49453 ovis aries
78	5	33.3	402	1	CPXE_STRGO	P18327 streptomyces
79	5	33.3	408	1	IYD_CAEEL	P34275 caenorhabdl
80	5	33.3	413	1	GAT1_HUMAN	P15976 homo sapien
81	5	33.3	421	1	SIAL_CHICK	Q92182 gallus gall
82	5	33.3	421	1	P2X5_HUMAN	Q93086 homo sapien
83	5	33.3	422	1	GAS7_RAT	Q55148 rattus norv
84	5	33.3	424	1	PAC3_HUMAN	Q9U566 homo sapien
85	5	33.3	424	1	PAC3_MOUSE	Q99138 mus musculu
86	5	33.3	425	1	POU1_BRAE	P31366 brachydanio
87	5	33.3	435	1	DPN_DROME	Q26263 drosophila
88	5	33.3	439	1	NTT4_HUMAN	Q9H1v8 homo sapien
89	5	33.3	439	1	RECA_ARATH	Q93199 arabidopsis
90	5	33.3	443	1	ZP23_BRAE	P79745 brachydanio
91	5	33.3	455	1	VNS1_BMDNV	P05540 bombyx dens
92	5	33.3	457	1	GBAL_DROME	P25157 drosophila
93	5	33.3	465	1	HXK4_HUMAN	P35557 homo sapien
94	5	33.3	465	1	HXK4_MOUSE	P52792 mus musculu
95	5	33.3	465	1	HXK4_RAT	P17712 rattus norv
96	5	33.3	471	1	PDM2_DROVI	Q24705 drosophila
97	5	33.3	471	1	SYE_ECO57	Q8XW02 escherichia
98	5	33.3	471	1	SYE_ECOLI	P04805 escherichia
99	5	33.3	471	1	SYE_RHOSH	Q9Zfz3 rhodobacter
100	5	33.3	471	1	SYE_SALTY	Q8Zfz9 salmonella
101	5	33.3	471	1	YHJ6_YEAST	P38770 saccharomyc
102	5	33.3	480	1	SYE_HAETN	P43818 haemophilus
103	5	33.3	480	1	SYE_PASMU	P57806 pasteurella
104	5	33.3	485	1	SYE_RHIME	P15780 rhizobium m
105	5	33.3	498	1	PDM2_DROME	P31369 drosophila
106	5	33.3	505	1	SCRV_KLEPN	P27218 klebsiella

THIS PAGE BLANK (USPTO)

RESULT 197

H71081

hypothetical protein PH0918 - Pyrococcus horikoshii

C:Species: Pyrococcus horikoshii

C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 28-Jul-2000

C:Accession: H71081

R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Seki, M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kishida, N.; Oguchi, DNA Res. 5, 55-76, 1998

A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic

A:Reference number: A71000; M01D:98344137; PMID:9679194

A:Accession: H71081

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-377 <RAM>

A:Cross-references: GB:AP000004; NID:g3236131; PIDN:BA030014.1; PID:g3257331

A:Experimental source: strain OT3

A:Note: this accession replaces an interim accession for a sequence replaced by GenBank

C:Genetics:

A:Gene: PH0918

C:Superfamily: Pyrococcus horikoshii hypothetical protein PH0918

Query Match

Best Local Similarity 100.0%; Score 5; DB 2; Length 377;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11

DB 57 RSTPE 61

RESULT 198

JC7535

chitinase (EC 3.2.1.14) 35 - Streptomyces thermoviolaceus

C:Species: Streptomyces thermoviolaceus

C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001

C:Accession: JC7535

R:Tsujiyama, H.; Okamoto, T.; Hatanu, N.; Miyamoto, K.; Watanabe, T.; Mitsunuma, M.; Inano, Biosci. Biotechnol. Biochem. 64, 2445-2453, 2000

A:Title: Family 19 chitinases from Streptomyces thermoviolaceus OPC-520: Molecular cloning

A:Reference number: JC7535; M01D:21056907; PMID:11193414

A:Accession: JC7535

A:Molecule type: DNA

A:Residues: 1-377 <TSU>

A:Cross-references: DDBJ:AB016842

A:Experimental source: strain OPC-520

C:Comment: This enzyme, a member of the family 19 chitinases, is involved in chitin degradation

C:Keywords: This enzyme, a member of the family 19 chitinases, is involved in chitin degradation

A:Gene: ch135

C:Keywords: glycosidase; hydrolase

Query Match

Best Local Similarity 100.0%; Score 5; DB 2; Length 377;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 PHRST 9

DB 6 PHRST 10

RESULT 199

S46711

hypothetical protein YHR063c - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein H8025.16

C:Species: Saccharomyces cerevisiae

C:Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 19-Apr-2002

C:Accession: S46711

R:Latreille, P.

submitted to the EMBL Data Library, May 1994

A:Description: The sequence of S. cerevisiae cosmid 8025.

A:Reference number: S46696

A:Accession: S46711

A:Molecule type: DNA

A:Residues: 1-379 <LAT>

A:Cross-references: EMBL:U00061; NID:g487943; PIDN:AMB68390.1; PID:g487955; MIPS:YHR0

C:Genetics:

A:Cross-references: SGD:S0001105

A:Map position: 8R

C:Keywords: transmembrane protein

Query Match

Best Local Similarity 100.0%; Score 5; DB 2; Length 379;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 PHRST 9

DB 4 PHRST 8

RESULT 200

TVMSF

transforming protein fos - mouse

C:Species: Mus musculus (house mouse)

C:Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 24-Sep-1999

C:Accession: A01343

R:Van Beveren, C.; van Straaten, F.; Curran, T.; Mueller, R.; Verma, I.M.

Cell 32, 1241-1255, 1983

A:Title: Analysis of FBJ-MuSV provirus and c-fos (mouse) gene reveals that viral and

A:Reference number: A03986; M01D:83180421; PMID:6301687

A:Accession: A01343

A:Molecule type: DNA

A:Residues: 1-380 <VAN>

A:Cross-references: GB:V00727; NID:g50399; PIDN:CAA24105.1; PID:g50400

C:Genetics:

A:Gene: c-fos

A:Introns: 47/3; 131/3; 167/3

C:Superfamily: fos transforming protein; fos/jun DNA-binding domain homology

C:Keywords: DNA binding; leucine zipper

F:132-172/Domain: fos/jun DNA-binding domain homology <FJD>

F:165-193/Region: leucine zipper motif

Query Match

Best Local Similarity 100.0%; Score 5; DB 1; Length 380;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPE 12

DB 231 STPE 235

Search completed: March 10, 2003, 14:30:46
Job time : 59 secs


```

RESULT 192
AB0407
Probable zinc-binding dehydrogenase YP03352 [Imported] - Yersinia pestis (strain C092)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
C:Accession: AB0407
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tlball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
ll, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AB0407
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-371 <KUR>
A:Cross-references: GB:AL590842; PIDN:CA092582.1; PID:q15981279; GSPDB:GN00175
C:Genetics:
A:Gene: YP03352
C:Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology

Query Match
Best Local Similarity 33.3%; Score 5; DB 2; Length 371;
Pred. No. 3,4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGP 5
DB 314 SHLGP 318

RESULT 193
T27643
Hypothetical protein ZK1010.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T27643
R:Gardner, A.
Submitted to the EMBL Data Library, November 1996
A:Reference number: Z20398
A:Accession: T27643
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-371 <WIL>
A:Cross-references: EMBL:Z82083; PIDN:CAB04972.1; GSPDB:GN00021; CESP:ZK1010.6
A:Experimental source: clone ZK1010
C:Genetics:
A:Gene: CESP:ZK1010.6
A:Map position: 3
A:Introns: 36/1; 71/1; 111/1; 175/2; 197/3; 237/1; 270/2; 308/3

Query Match
Best Local Similarity 33.3%; Score 5; DB 2; Length 371;
Pred. No. 3,4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRA 14
DB 79 PESRA 83

RESULT 194
AD0719
Probable bacteriophage protein STY1893 [Imported] - Salmonella enterica subsp. enterica
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
C:Accession: AD0719
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Garra, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; PMID:11677608

```

```

A:Accession: AD0719
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-374 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD02125.1; PID:q16502959; GSPDB:GN00176
C:Genetics:
A:Gene: STY1893

Query Match
Best Local Similarity 33.3%; Score 5; DB 2; Length 374;
Pred. No. 3,4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRA 14
DB 48 PESRA 52

RESULT 195
T23705
Hypothetical protein M04C7.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T23705
R:Kershaw, J.
Submitted to the EMBL Data Library, November 1996
A:Reference number: Z19786
A:Accession: T23705
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-375 <WIL>
A:Cross-references: EMBL:Z83117; PIDN:CAB05570.1; GSPDB:GN00019; CESP:M04C7.1
A:Experimental source: clone M04C7
C:Genetics:
A:Gene: CESP:M04C7.1
A:Map position: 1
A:Introns: 55/2; 102/3; 156/2; 199/2; 218/2; 261/3; 313/1
C:Superfamily: GTP-binding regulatory protein Gs alpha chain

Query Match
Best Local Similarity 33.3%; Score 5; DB 2; Length 375;
Pred. No. 3,4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 SPES 12
DB 6 SPES 10

RESULT 196
S21302
Succinate dehydrogenase (EC 1.3.99.1) flavoprotein - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 20-May-1994 #sequence_revision 10-Nov-1995 #text_change 11-Jun-1999
C:Accession: S21302
R:Malcovati, M.; Marchetti, L.; Zanello, E.; Tencin, M.L.; Simonc, T.; Benatti, L.
Submitted to the EMBL Data Library, July 1990
A:Description: Cloning of the flavoprotein subunit of human succinate dehydrogenase.
A:Reference number: S21302
A:Accession: S21302
A:Molecule type: mRNA
A:Residues: 1-377 <MAL>
A:Cross-references: EMBL:X53943; NID:q36642; PIDN:CAA37886.1; PID:q36643
C:Superfamily: fumarate reductase flavoprotein; 3-oxosteroid 1-dehydrogenase homology
C:Keywords: flavoprotein; mitochondrion; oxidoreductase
F:149-239/Domains: 3-oxosteroid 1-dehydrogenase homology <OXD>

Query Match
Best Local Similarity 33.3%; Score 5; DB 2; Length 377;
Pred. No. 3,4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRAA 15
DB 369 ESRAA 373

```

A:Title: Planarian homeobox genes: cloning, sequence analysis, and expression.
A:Reference number: A41151; MWID:91334461; PMID:1714599
A:Accession: B41151
A:Molecule type: mRNA
A:Residues: 1-363 <GA2>
A:Cross-references: GB:X56500; NID:99157; PIDN:CAA39855.1; PID:99158
A:Experimental source: intact and regenerating planarians
C:Genetics:
A:Introns: 75/1; 176/3
C:Superfamily: homeotic protein Hox B2; homeobox homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:134-190/Domain: homeobox homology <HOX>

Query Match 33.3%; Score 5; DB 2; Length 363;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
|||||
DB 231 RSTPE 235

RESULT 188
AD3557
acriflavin resistance protein e [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: AD3557
R:DeVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Petra, G.; Mujer, C.; Los, T.; Ivanova, .; Mazur, M.; Goldsman, E.; Selkov, E.; Elser, P.H.; Hagius, S.; O'Callaghan, D.; Letess, Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:11756688
A:Accession: AD3557
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-364 <KUR>
A:Cross-references: GB:AE008918; PIDN:AAL53623.1; PID:g17984538; GSPDB:GN00191
C:Experimental source: strain 16M
C:Genetics:
A:Gene: BME110381
A:Map position: 11

Query Match 33.3%; Score 5; DB 2; Length 364;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
|||||
DB 281 RSTPE 285

RESULT 189
A34424
CD44 membrane glycoprotein precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 15-Jun-1990 #sequence_revision 15-Jun-1990 #text_change 01-Dec-2000
C:Accession: A34424; A34907
R:Notenbury, C.; Rees, G.; St. John, T
Proc. Natl. Acad. Sci. U.S.A. 86, 8521-8525, 1989
A:Title: Isolation of mouse CD44 cDNA: structural features are distinct from the primate
A:Reference number: A34424; MWID:90046829; PMID:2682651
A:Accession: A34424
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-365 <NOT>
A:Cross-references: GB:M27130; NID:9192530; PIDN:AAA37407.1; PID:9309161
R:Wolfe, E.J.; Gause, W.C.; Pelfrey, C.M.; Holland, S.M.; Steinberg, A.D.; August, J.T.
J. Biol. Chem. 265, 341-347, 1990
A:Title: The cDNA sequence of mouse Pgp-1 and homology to human CD44 cell surface antigen
A:Reference number: A34907; MWID:90094420; PMID:2403559
A:Accession: A34907
A:Status: preliminary

A:Molecule type: mRNA
A:Residues: 8-195; 'G', 197-365 <MOL>
A:Cross-references: GB:J05163; NID:9200334; PIDN:AAA39923.1; PID:9200335
C:Superfamily: human cell adhesion protein CD44
C:Keywords: cell adhesion; glycoprotein; membrane protein

Query Match 33.3%; Score 5; DB 2; Length 365;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPE 12
|||||
DB 192 STPE 196

RESULT 190
S51363
actin modulator protein ENAM - earthworm (Lumbricus terrestris)
C:Species: Lumbricus terrestris (common earthworm)
C:Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 26-Aug-1999
C:Accession: S51363; S49320
R:Clebing, T.; Hinssen, H.; D'Haese, J.
Eur. J. Biochem. 225, 773-779, 1994
A:Title: The complete sequence of a 40-kDa actin-modulating protein from the earthworm
A:Reference number: S51363; MWID:95045553; PMID:7957213
A:Accession: S51363
A:Molecule type: mRNA
A:Residues: 1-366 <GIE>
A:Cross-references: EMBL:232528; NID:9551451; PID:9551452
A:Note: part of this sequence was confirmed by protein sequencing
C:Superfamily: severin; gelsolin repeat homology
F:17-366/Domain: gelsolin repeat homology <GEL>

Query Match 33.3%; Score 5; DB 2; Length 366;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPE 12
|||||
DB 245 STPE 249

RESULT 191
G83180
probable FMN oxidoreductase PA3723 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: G83180
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; I
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; I
., Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A:Reference number: AB2950; MWID:20437337; PMID:10984043
A:Accession: G83180
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-368 <STO>
A:Cross-references: GB:AE004791; GB:AE004091; NID:99949882; PIDN:AAG07110.1; GSPDB:GN
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA3723
C:Superfamily: NADPH dehydrogenase chain OYE2

Query Match 33.3%; Score 5; DB 2; Length 368;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 HRSTP 10
|||||
DB 356 HRSTP 360

A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: H70882
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-356 <COL>
A:Cross-references: GB:AI008967; GB:AI123456; NID:93261491; PIDN:CAA15572.1; PID:9262429
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv2777c
C:Superfamily: Mycobacterium tuberculosis hypothetical protein Rv2777c

Query Match 33.3%; Score 5; DB 2; Length 356;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRA 15
DB 66 ESRA 70

RESULT 183
C98171
hypothetical protein AGR_L_636 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C:Species: Agrobacterium tumefaciens
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C:Accession: C98171
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Mollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; PMID:11743194
A:Accession: C98171
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-356 <KJB>
A:Cross-references: GB:AE007870; PIDN:AAK8893.1; PID:915158664; GSPDB:GN00170
C:Genetics:
A:Gene: AGR_L_636
A:Map position: linear chromosome

Query Match 33.3%; Score 5; DB 2; Length 356;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESR 13
DB 298 TPESR 302

RESULT 184
I38920
prostaglandin E2 receptor - human
C:Species: Homo sapiens (man)
C:Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 23-Jul-1999
C:Accession: I38920
R:Regan, J.W.; Bailey, T.J.; Pepperl, D.J.; Pierce, K.L.; Bogardus, A.M.; Donello, J.E.;
Mol. Pharmacol. 46, 213-220, 1994
A:Title: Cloning of a novel human prostaglandin receptor with characteristics of the pha
A:Reference number: I38920; MUID:94359483; PMID:8078484
A:Accession: I38920
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-358 <RES>
A:Cross-references: EMBL:U19487; NID:9639719; PIDN:AA61681.1; PID:9632650
C:Superfamily: prostaglandin E receptor EPI

Query Match 33.3%; Score 5; DB 2; Length 358;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRA 14
DB 104 PESRA 108

RESULT 185
S51312
EP2 prostaglandin receptor - human
C:Species: Homo sapiens (man)
C:Date: 07-May-1995 #sequence_revision 01-Sep-1995 #text_change 23-Jul-1999
C:Accession: S51312
R:Oakley, C.J.
submitted to the EMBL Data Library, January 1995
A:Reference number: S51312
A:Accession: S51312
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-358 <OAK>
A:Cross-references: EMBL:X83868; NID:9633205; PIDN:CAA58749.1; PID:9633206
C:Superfamily: prostaglandin E receptor EPI

Query Match 33.3%; Score 5; DB 2; Length 358;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRA 14
DB 104 PESRA 108

RESULT 186
G87706
hypothetical protein CC3689 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: G87706
R:Nierman, W.C.; Felblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: G87706
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-361 <STO>
A:Cross-references: GB:AE005673; NID:913425451; PIDN:AAK25651.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC3689

Query Match 33.3%; Score 5; DB 2; Length 361;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRA 15
DB 183 ESRA 187

RESULT 187
S33702
homeotic protein Dth-2 - planarian (Dugesia tigrina)
C:Species: Dugesia tigrina
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 19-Jul-2002
C:Accession: S33702; B41151
R:Garcia-Fernandez, J.; Baguna, J.; Salo, E.
Development 118, 241-253, 1993
A:Title: Genomic organization and expression of the planarian homeobox genes Dth-1 an
A:Reference number: S33701; MUID:93387216; PMID:8104142
A:Accession: S33702
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-363 <GAR>
A:Cross-references: EMBL:X69202; NID:9312390; PIDN:CAA9140.1; PID:9312391
R:Garcia-Fernandez, J.; Baguna, J.; Salo, E.
Proc. Natl. Acad. Sci. U.S.A. 88, 7338-7342, 1991

Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTRP 11
|||||
Db 97 RSTRP 101

RESULT 178
G85638
hypothetical protein 21444 [Imported] - Escherichia coli (strain O157:H7, substrain EDL5
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: G85638
R:Perera, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Illier, L.; Gotlibeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: AB5480; MUID:21074935; PMID:1120551
A:Accession: G85638
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-348 <STO>
A:Cross-references: GB:AE005174; NID:g12514297; PIDN:AG655571.1; GSPDB:GN00145; UNCP:214
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: 21444

Query Match 33.3%; Score 5; DB 2; Length 348;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRAA 15
|||||
Db 160 ESRAA 164

RESULT 179
T24015
hypothetical protein R07B7.9 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
C:Accession: T24015
R:Harris, B.
submitted to the EMBL Data Library, July 1996
A:Reference number: 219830
A:Accession: T24015
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-349 <WIL>
A:Cross-references: EMBL:Z75955; PIDN:CAB00117.1; GSPDB:GN00023; CESP:R07B7.9
A:Experimental source: clone R07B7
C:Genetics:
A:Gene: CESP:R07B7.9
A:Map position: 5
A:Introns: 4/1; 55/3; 92/1; 148/3; 224/2; 278/3; 336/3
C:Superfamily: Caenorhabditis elegans hypothetical protein R07B7.8

Query Match 33.3%; Score 5; DB 2; Length 349;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESR 13
|||||
Db 229 TPESR 233

RESULT 180
D96761
unknown protein [Imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: D96761

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luos, J.S.; Malt, R.; Matzla
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: AB6141; MUID:21016719; PMID:11130712
A:Accession: D96761
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-351 <STO>
A:Cross-references: GB:AE005173; NID:g11120789; PIDN:AG30969.1; GSPDB:GN00141
C:Genetics:
A:Gene: T9L24.34
A:Map position: 1

Query Match 33.3%; Score 5; DB 2; Length 351;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRAA 15
|||||
Db 277 ESRAA 281

RESULT 181
F83190
Probable UDP-3-O-[3-hydroxymyristoyl] glucosamine N-acetyltransferase (EC 2.3.1.-) PA36
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 02-Feb-2001
C:Accession: F83190
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lathig, K.; L
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: F83190
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-353 <STO>
A:Cross-references: GB:AE004784; GB:AE004091; NID:g9949799; PIDN:AG07034.1; GSPDB:GN
A:Experimental source: strain PA01
C:Genetics:
A:Gene: lpxD; PA3646
C:Superfamily: UDP-3-O-[3-hydroxymyristoyl] glucosamine N-acetyltransferase
C:Keywords: acyltransferase

Query Match 33.3%; Score 5; DB 2; Length 353;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRA 14
|||||
Db 55 PESRA 59

RESULT 182
H70882
hypothetical protein RV2777c - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: H70882
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
; Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno

RESULT 173
AB3043
oxidoreductase Atu3965 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: AB3043
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AB3043
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-338 <KUR>
A:Cross-references: GB:AE008689; PIDN:AL44767.1; PID:g17742404; GSPDB:GN00187
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu3965
A:Map position: linear chromosome

Query Match 33.3%; Score 5; DB 2; Length 338;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPHR 7
|||||
DB 155 LGPHR 159

RESULT 174
BB3757
partial probable outer membrane channel protein 22506 [imported] - Escherichia coli (str
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: BB3757
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobeck, E.J.; Davis, N.W.; Llim, A.; Diallanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: AB5480; MUID:21074935; PMID:11206551
A:Accession: BB3757
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-342 <STO>
A:Cross-references: GB:AE005174; NID:g12515491; PIDN:AGS6518.1; GSPDB:GN00145; UWGP:Z25
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: 22506

Query Match 33.3%; Score 5; DB 2; Length 342;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STPES 12
|||||
DB 24 STPES 28

RESULT 175
AB3116
transcription regulator, lact family Atu4550 [imported] - Agrobacterium tumefaciens (str
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: AB3116
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AB3116
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-342 <KUR>
A:Cross-references: GB:AE008689; PIDN:AL45344.1; PID:g17743037; GSPDB:GN00187
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu4550
A:Map position: linear chromosome

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AB3116
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-342 <KUR>
A:Cross-references: GB:AE008689; PIDN:AL45344.1; PID:g17743037; GSPDB:GN00187
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu4550
A:Map position: linear chromosome

Query Match 33.3%; Score 5; DB 2; Length 342;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 TPESR 13
|||||
DB 284 TPESR 288

RESULT 176
C75139
bifunctional short chain isoprenyl diphosphate synthase (ldsA) PAB2389 - Pyrococcus a
C:Species: Pyrococcus abyssi
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C:Accession: C75139
R:anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome s
A:Reference number: A75001
A:Accession: C75139
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-346 <KAN>
A:Cross-references: GB:AJ248285; GB:AL096836; NID:g5458067; PIDN:CAB49828.1; PID:g545
C:Genetics:
A:Gene: PAB2389
C:Superfamily: prenyl transferase A

Query Match 33.3%; Score 5; DB 2; Length 346;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRA 14
|||||
DB 326 PESRA 330

RESULT 177
T47411
hypothetical protein T28A8.40 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T47411
R:Punelle, B.; Boutry, M.; Goffeau, A.; Mewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.
submitted to the Protein Sequence Database, March 2000
A:Reference number: 224466
A:Accession: T47411
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-346 <PUR>
A:Cross-references: EMBL:AL162691
A:Experimental source: cultivar Columbia; BAC clone T28A8
C:Genetics:
A:Map position: 3
A:Introns: 93/3
A:Note: T28A8.40

Query Match 33.3%; Score 5; DB 2; Length 346;

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hlizar, L.
Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maltl, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sekano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shin, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
A:Reference number: A86141; MIMD:21016719; PMID:11130712
A:Accession: B86429
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-335 <STO>
A:Cross-references: GB:AE005172; NID:g6634766; PIDN:AAF19746.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 33.3%; Score 5; DB 2; Length 335;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRA 14
|||||
DB 105 PESRA 109

RESULT 169
T44988
oxido-reductase [imported] - Haloflex volcani megaplasmid pHV3
C:Species: Haloflex volcani
C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 17-Mar-2000
C:Accession: T44988
R:Farhani, R.; Imbeault, J.C.; St Jean, A.; Chan, C.C.Y.; Allard, G.; Charlebois, R.L.
submitted to the EMBL Data Library, March 1997
A:Description: Hereditary instability of the megaplasmid pHV3, and filamentation in the
A:Reference number: Z22886
A:Accession: T44988
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-336 <FAR>
A:Cross-references: EMBL:U95374; PIDN:AA071807.1
A:Experimental source: strain D52
C:Genetics:
A:Map position: megaplasmid pHV3
A:Genome: plasmid
A:Note: expressed during exponential growth
C:Superfamily: fission yeast pyridoxine 4-dehydrogenase

Query Match 33.3%; Score 5; DB 2; Length 336;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRA 15
|||||
DB 243 ESRA 247

RESULT 170
E83059
ketol-acid reductoisomerase PA4694 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: E83059
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bz
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lam,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
A:Reference number: A82950; MIMD:20437337; PMID:10984043
A:Accession: E83059
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-338 <STO>
A:Cross-references: GB:AE004883; GB:AE004091; NID:g9950947; PIDN:AA08080.1; GSPDB:G
A:Experimental source: strain PA01
C:Genetics:
A:Gene: llyC; PA4694
C:Superfamily: Methanococcus ketol-acid reductoisomerase; ketol-acid reductoisomerase

Query Match 33.3%; Score 5; DB 2; Length 338;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRA 15
|||||
DB 269 ESRA 273

RESULT 171
A83835
NADH oxidase BH1481 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: A83835
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* a
A:Reference number: A83650; MIMD:20512582; PMID:11058132
A:Accession: A83835
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-338 <STO>
A:Cross-references: GB:AP001512; GB:BA000004; NID:g10174030; PIDN:BA05200.1; GSPDB:G
C:Genetics:
A:Gene: BH1481
C:Superfamily: NADPH dehydrogenase chain OYE2

Query Match 33.3%; Score 5; DB 2; Length 338;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRA 15
|||||
DB 46 ESRA 50

RESULT 172
C98242
hypothetical protein AGR_L1780 [imported] - Agrobacterium tumefaciens (strain C58, C
C:Species: Agrobacterium tumefaciens
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C:Accession: C98242
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Gold
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Marzels,
Science 294, 2323-2328, 2001
A:Title: Genome sequence of the plant pathogen and biotechnology agent *Agrobacterium*
A:Reference number: A97359; PMID:11743194
A:Accession: C98242
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-338 <GR>
A:Cross-references: GB:AE007870; PIDN:AAK89461.1; PID:g15159327; GSPDB:GN00170
C:Genetics:
A:Gene: AGR_L1780
A:Map position: linear chromosome

Query Match 33.3%; Score 5; DB 2; Length 338;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPHR 7
|||||
DB 155 LGPHR 159

A.Molecule type: DNA
 A.Residues: 1-332 <AOF>
 A.Cross-references: GB:AE000719; NID:g2983517; PIDN:AA07097.1; PID:g2983521; GB:AE00065
 A.Experimental source: strain VRS
 C.Genetics:
 A.Gene: hlyE
 C:Superfamily: hydrogenase expression/formation protein hlyE

Query Match 33.3%; Score 5; DB 2; Length 332;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 7 RSTPE 11
 |||||
 DB 286 RSTPE 290

RESULT 164
 157032
 gene Tlx-1 protein - mouse
 C:Species: Mus sp. (mouse)
 C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 24-Sep-1999
 C:Accession: 157032
 R:Ray, K.; Tang, S.; Dube, I.D.; Kamel-Reid, S.; Bryce, D.M.; Breitman, M.L.
 Mech. Dev. 44, 51-64, 1993
 A>Title: Characterization and developmental expression of Tlx-1, the murine homolog of H
 A:Reference number: 157032; MUID:94206842; PMID:7908826
 A:Accession: 157032
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A.Molecule type: DNA
 A.Residues: 1-332 <RES>
 A.Cross-references: GB:S70632; NID:g546379; PIDN:AB30542.1; PID:g546380
 C.Genetics:
 A.Gene: Tlx-1
 A:introns: 192/1; 259/2
 C:Superfamily: unassigned homeobox proteins; homeobox homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F:204-260/Domain: homeobox homology <HOX>

Query Match 33.3%; Score 5; DB 2; Length 332;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 HLGPH 6
 |||||
 DB 3 HLGPH 7

RESULT 165
 AG0809
 Probable membrane protein STY2661 [imported] - Salmonella enterica subsp. enterica serov
 C:Species: Salmonella enterica subsp. enterica serovar Typh
 A>Note: This species has also been called Salmonella typh
 C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
 C:Accession: AG0809
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher,
 th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
 S.; Moule, S.; O'Garra, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
 A>Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
 A:Reference number: AB0502; PMID:11677608
 A:Accession: AG0809
 A:Status: preliminary
 A.Molecule type: DNA
 A.Residues: 1-332 <PAR>
 A.Cross-references: GB:AL513382; PIDN:CAD07657.1; PID:g16503644; GSPDB:GN00176
 C.Genetics:
 A.Gene: STY2661

Query Match 33.3%; Score 5; DB 2; Length 332;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 11 ESRAA 15
 |||||
 DB 326 ESRAA 330

RESULT 166
 C22735
 hypothetical nox2 protein - Emeritella nidulans mitochondrion
 C:Species: mitochondrion Emeritella nidulans, Aspergillus nidulans
 C>Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 21-Jul-2000
 C:Accession: C22735
 R:Marings, R.B.; Brown, T.A.; Ray, J.A.; Scazzocchio, C.; Davies, R.W.
 EMBO J. 3, 2121-2128, 1984
 A>Title: Three variant introns of the same general class in the mitochondrial gene fo
 A:Reference number: AB0991; MUID:85027165; PMID:6092056
 A:Accession: C22735
 A.Molecule type: DNA
 A.Residues: 1-334 <MAR>
 A.Cross-references: GB:X00790; NID:g12690; PIDN:CA25371.1; PID:g134387
 C.Genetics:
 A.Gene: mitochondrion
 A:Genetic code: SGC3
 C:Superfamily: COI intron 7a protein
 C:Keywords: mitochondrion

Query Match 33.3%; Score 5; DB 2; Length 334;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 7 RSTPE 11
 |||||
 DB 248 RSTPE 252

RESULT 167
 AB3342
 conserved hypothetical protein PA2430 [imported] - Pseudomonas aeruginosa (strain PAO
 C:Species: Pseudomonas aeruginosa
 C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: AB3342
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; L
 .; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A>Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
 A:Reference number: AB2950; MUID:20437337; PMID:10984043
 A:Accession: AB3342
 A:Status: preliminary
 A.Molecule type: DNA
 A.Residues: 1-334 <STO>
 A.Cross-references: GB:AE004670; GB:AE004091; NID:g9948470; PIDN:AA05818.1; GSPDB:GN
 A:Experimental source: strain PA01
 C.Genetics:
 A.Gene: PA2430
 C:Superfamily: Escherichia coli hypothetical protein b1644

Query Match 33.3%; Score 5; DB 2; Length 334;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 11 ESRAA 15
 |||||
 DB 102 ESRAA 106

RESULT 168
 B86429
 P26G16.5 protein - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Nov-2001
 C:Accession: B86429
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon

Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.: Li, J.H.; Li, Y.; Lin, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzall,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shin, P.; Southwick, A.M.; Sun, H.; Tallon,
 Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: F96522
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-326 <STO>
 A:Cross-references: GB:AE005173; NID:95733881; PIDN:AAD49769.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: FL1A17.17
 A:Map position: 1

Query Match 33.3%; Score 5; DB 2; Length 326;
 Best Local Similarity 100.0%; Pred. No. 3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHGP 5
 |||||
 DB 145 SHGP 149

RESULT 160
 A84291
 ornithine cyclodeaminase [imported] - Halobacterium sp. NRC-1
 C:Species: Halobacterium sp. NRC-1
 C:Date: 02-Feb-2001 #sequence-revision 02-Feb-2001 #text-change 02-Feb-2001
 C:Accession: A84291
 R:Ng, W.V.; Kennedy, S.P.; Mahatras, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.
 ; Leithauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jadic
 Jung, K.H.; Alam, M.; Freitas, T.
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
 A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li
 A:Title: Genome sequence of Halobacterium species NRC-1.
 A:Reference number: A84160; MUID:20504483; PMID:11016950
 A:Accession: A84291
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-328 <STO>
 A:Cross-references: GB:AE004437; NID:910580873; PIDN:AAG19693.1; GSPDB:GN00138
 C:Genetics:
 A:Gene: ocd2

Query Match 33.3%; Score 5; DB 2; Length 328;
 Best Local Similarity 100.0%; Pred. No. 3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESR 13
 |||||
 DB 154 TPESR 158

RESULT 161
 A40855
 homeotic protein Hox 11 - human
 N:Alternate names: tcl-3 proto-oncogene
 C:Species: Homo sapiens (man)
 C:Date: 28-Feb-1992 #sequence-revision 28-Feb-1992 #text-change 24-Sep-1999
 C:Accession: A40855; S17311; A41224
 R:Altano, M.; Roberts, C.W.M.; Minden, M.; Crist, W.M.; Korsmeyer, S.J.
 Science 253, 79-82, 1991
 A:Title: Deregulation of a homeobox gene, HOX11, by the t(10;14) in T cell leukemia.
 A:Reference number: A40855; MUID:91289163; PMID:1676542
 A:Accession: A40855
 A:Molecule type: mRNA
 A:Residues: 1-330 <HAT>
 A:Cross-references: GB:S38742; NID:9232582; PIDN:AAB19293.1; PID:9232583
 R:Lu, M.; Gong, Z.; Shen, W.; Ho, A.D.
 EMBO J. 10, 2905-2910, 1991

A:Title: The tcl-3 proto-oncogene altered by chromosomal translocation in T-cell leuk
 A:Reference number: S17311; MUID:92007734; PMID:1717256
 A:Accession: S17311
 A:Molecule type: mRNA
 A:Residues: 1-275, 'L', 277-330 <LUX>
 A:Cross-references: EMBL:M62626; NID:9339199; PIDN:AAA36719.1; PID:9387683
 R:Kennedy, M.A.; Gonzalez-Sarmiento, R.; Kees, U.R.; Lampert, F.; Dear, N.; Boehm, T
 Proc. Natl. Acad. Sci. U.S.A. 88, 8900-8904, 1991
 A:Title: HOX11, a homeobox-containing T-cell oncogene on human chromosome 10q24.
 A:Reference number: A41224; MUID:92020958; PMID:1681546
 A:Accession: A41224
 A:Molecule type: DNA
 A:Residues: 1-189, 'VALSPFTVTRR', 190-275, 'L', 277-330 <KEN>
 A:Cross-references: GB:M75952
 A:Note: the authors suggest a long form with twelve additional amino acids from poss
 C:Genetics:
 A:Gene: GDB:HOX11
 A:Cross-references: GDB:119607; OMIM:186770
 A:Map position: 10q24-10q24
 C:Superfamily: unassigned homeobox proteins; homeobox homology
 C:Keywords: DNA binding; homeobox; leukemia; nucleus; proto-oncogene; transcription
 F:202-258/Domain: homeobox homology <HOX>

Query Match 33.3%; Score 5; DB 2; Length 330;
 Best Local Similarity 100.0%; Pred. No. 3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HLGP 6
 |||||
 DB 3 HLGP 7

RESULT 162
 S44743
 C0205.1 protein - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 16-Sep-1994 #sequence-revision 12-May-1995 #text-change 19-May-2000
 C:Accession: S44743
 R:Du, Z.
 submitted to the EMBL Data Library, May 1993
 A:Description: Sequence of the C. elegans cosmid C0205.
 A:Reference number: S44613
 A:Accession: S44743
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-332 <DUZ>
 A:Cross-references: EMBL:L16622; NID:9289603; PID:9289604
 C:Genetics:
 A:introns: 48/1; 165/3; 283/3
 C:Superfamily: acyl-CoA dehydrogenase

Query Match 33.3%; Score 5; DB 2; Length 332;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 HRSTP 10
 |||||
 DB 141 HRSTP 145

RESULT 163
 A70388
 hydroxase expression/formation protein - Aquifex aeolicus
 C:Species: Aquifex aeolicus
 C:Date: 08-May-1998 #sequence-revision 08-May-1998 #text-change 24-Sep-1999
 C:Accession: A70388
 R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
 V.
 Nature 392, 353-358, 1998
 A:Title: The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.
 A:Reference number: A70300; MUID:98196666; PMID:9537320
 A:Accession: A70388
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown

AB0330
phospholipase A [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C:Accession: AB0330
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tiltball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Hill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AB0330
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-320 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC92945.1; PID:g15980684; GSPDB:GN00175
C:Genetics:
A:Gene: ypla

Query Match 33.3%; Score 5; DB 2; Length 320;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 HRSTP 10
|||||
Db 47 HRSTP 51

RESULT 155
A43681
Immediate-early protein PE-38 - Autographa californica nuclear polyhedrosis virus
C:Species: Autographa californica nuclear polyhedrosis virus, ACNMPV
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
C:Accession: A43681
R:Krappe, R.; Knebel-Moersdorf, D.
J. Virol. 65, 805-812, 1991
A:Title: Identification of the very early transcribed baculovirus gene PE-38.
A:Reference number: A43681; MUID:91101290; PMID:1987375
A:Accession: A43681
A:Molecule type: DNA
A:Residues: 1-321 <KRA>
A:Cross-references: GB:M62468; NID:g332470; PIDN:AAA6733.1; PID:g332471
C:Superfamily: Orgyia pseudotsugata nuclear polyhedrosis virus PE-38 protein
C:Keywords: DNA binding; Immediate-early protein

Query Match 33.3%; Score 5; DB 1; Length 321;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 HRSTP 10
|||||
Db 9 HRSTP 13

RESULT 156
C72869
Hypothetical protein - Autographa californica nuclear polyhedrosis virus
C:Species: Autographa californica nuclear polyhedrosis virus, ACNMPV
A:Note: dsDNA virus
C:Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 11-Jan-2000
C:Accession: C72869
R:AYres, M.D.; Howard, S.C.; Kuzio, J.; Lopez-Ferber, M.; Possee, R.D.
Virology 202, 586-605, 1994
A:Title: The complete DNA sequence of Autographa californica nuclear polyhedrosis virus.
A:Reference number: A72850; MUID:94303173; PMID:8030224
A:Accession: C72869
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-321 <AVR>
A:Cross-references: GB:L22858; NID:g510708; PIDN:AAA66783.1; PID:g559222
C:Genetics:
A:Gene: AC-PE38
C:Superfamily: Orgyia pseudotsugata nuclear polyhedrosis virus PE-38 protein

Query Match 33.3%; Score 5; DB 2; Length 321;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 HRSTP 10
|||||
Db 9 HRSTP 13

RESULT 157
S55640
Hypothetical protein 45 - equine herpesvirus 2
C:Species: equine herpesvirus 2
C:Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 08-Oct-1999
C:Accession: S55640
R:Telford, E.A.R.; Watson, M.S.; Alrd, H.C.; Perry, J.; Davison, A.J.
J. Mol. Biol. 249, 520-528, 1995
A:Title: The DNA sequence of equine herpesvirus 2.
A:Reference number: S55594; MUID:95302501; PMID:7783207
A:Accession: S55640
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-321 <TEL>
A:Cross-references: GB:U20824; NID:g695172; PIDN:AAC13833.1; PID:g695218
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1995

Query Match 33.3%; Score 5; DB 2; Length 321;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPE 11
|||||
Db 182 RSTPE 186

RESULT 158
T13005
Hypothetical protein T24C20.40 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 13-Aug-1989 #sequence_revision 13-Aug-1989 #text_change 22-Oct-1999
C:Accession: T13005
R:Choline, N.; Robert, C.; Brotlier, P.; Wincker, P.; Catolico, L.; Artiguenave, F.; submitted to the Protein Sequence Database, July 1999
A:Reference number: 217586
A:Accession: T13005
A:Molecule type: DNA
A:Residues: 1-323 <CHO>
A:Cross-references: EMBL:AL096856; GSPDB:GN00061; ATSP:T24C20.40
A:Experimental source: cultivar Columbia; BAC clone T24C20
C:Genetics:
A:Gene: ATSP:T24C20.40
A:Map position: 3
A:Introns: 12/1; 29/3; 55/3; 73/3; 111/1; 133/3; 159/1; 179/3

Query Match 33.3%; Score 5; DB 2; Length 323;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STPE 12
|||||
Db 250 STPE 254

RESULT 159
F96522
Hypothetical protein F1A17.17 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: F96522
R:Thellogis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, ansen, N.F.; Hughes, B.; Huizar, L.

C:Superfamily: probable pyruvate formate-lyase 2 activating enzyme; ferredoxin 2[AFe-4S]
C:Keywords: iron; metalloprotein; oxidoreductase
F:34,38,41/Binding site: iron (Cys) #status predicted

Query Match

33.3%; Score 5; DB 1; Length 308;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRA 14
|||||

DB 44 PESRA 48

RESULT 150

T41889

PE38 orf133 - Bombyx mori nuclear polyhedrosis virus (isolate T3)

C:Species: Bombyx mori nuclear polyhedrosis virus; BMSNPV

A:Variety: isolate T3

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000

C:Accession: T41889

R:Gomi, S.; Majima, K.; Maeda, S.

J. Gen. Virol. 80, 1323-1337, 1999

A:Title: Sequence analysis of the genome of Bombyx mori nucleopolyhedrovirus.

A:Reference number: Z2020; MUID:99281911; PMID:10355780

A:Accession: T41889

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-309 <RAM>

A:Cross-references: EMBL:U33180; NID:g3745835; PIDN:AACG3818.1; PID:g3745971

A:Experimental source: isolate T3

C:Genetics:

A:Note: pe38

C:Superfamily: Orgyia pseudotsugata nuclear polyhedrosis virus PE-38 protein

Query Match

33.3%; Score 5; DB 2; Length 309;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 HRSTP 10
|||||

DB 11 HRSTP 15

RESULT 151

T26531

hypothetical protein Y18D10A.9 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 26-May-2000

C:Accession: T26531

R:Haris, B.

submitted to the EMBL Data Library, December 1998

A:Reference number: Z20226

A:Accession: T26531

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-314 <WID>

A:Cross-references: EMBL:AL034393; PIDN:CAA2322.1; CESP:Y18D10A.9

A:Experimental source: clone Y18D10A

C:Genetics:

A:Gene: CESP:Y18D10A.9

A:introns: 51/1; 140/1; 169/3; 203/3; 257/2

C:Superfamily: unassigned WD repeat proteins; WD repeat homology

Query Match

33.3%; Score 5; DB 2; Length 314;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPEs 12
|||||

DB 263 STPEs 267

RESULT 152

F87260

WeeB/YagA/Cpsf family protein [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C:Accession: F87260

R:Nietman, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,

B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwin, M.L.; Haft, D.H.; Ko

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete genome sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: F87260

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-316 <STO>

A:Cross-references: GB:AE005673; NID:g13421198; PIDN:AAK22082.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC0095

Query Match

33.3%; Score 5; DB 2; Length 316;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRRA 15
|||||

DB 296 ESRRA 300

RESULT 153

F82672

Arp sulfurylase, small subunit XF1500 [imported] - Xylella fastidiosa (strain 9a5c)

C:Species: Xylella fastidiosa

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000

C:Accession: F82672

R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq

Nature 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: A82515; MUID:20365717; PMID:10910347

A:Note: for a complete list of authors see reference number A59328 below

A:Accession: F82672

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-317 <SIM>

A:Cross-references: GB:AE003849; NID:g9106531; PIDN:AAFP4309.1; GSPDB:GN

A:Experimental source: strain 9a5c

R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carr

as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kurame, E.E.; La

chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins

A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.

, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa

A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv

M.; Tsuchioka, M.H.; Vajlada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.

A:Reference number: A59328

A:Contents: annotation

C:Genetics:

A:Gene: XF1500

C:Superfamily: nucleation protein nodp

Query Match

33.3%; Score 5; DB 2; Length 317;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRRA 15
|||||

DB 274 ESRRA 278

RESULT 154

C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: F84329
R:Ng, W.V.; Kennedy, S.P.; Mahairs, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leitner, B.; Keller, K.; Cruik, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jaidic Jung, K.H.; Alam, M.; Freltas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: F84329
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-302 <STO>
A:Cross-references: GB:AE004437; NID:g10581234; PIDN:AAG2002.1; GSPDB:GN00138
C:Genetics:
A:Gene: VNC1779C

Query Match 33.3%; Score 5; DB 2; Length 302;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 11 ESRAA 15
Db 248 ESRAA 252

RESULT 146

S75481

POLYALIC acid transport protein kpsM - Synecchocystis sp. (strain PCC 6803)

N:Alternate names: protein Slr2107

C:Species: Synecchocystis sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999

C:Accession: S75481

R:Kenno, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;

O, K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda

DNA Res. 3, 109-116, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecchocystis

S.

A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S75481

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-302 <KAN>

A:Cross-references: EMBL:D90911; GB:AB001339; NID:g1653083; PIDN:BAA18042.1; PID:d101877

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Genetics:

A:Gene: kpsM

A:Start codon: GTG

Query Match 33.3%; Score 5; DB 2; Length 302;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 9 TPESR 13
Db 29 TPESR 33

RESULT 147

G64405

tetrahydromethanopterin S-methyltransferase (EC 2.1.1.86) chain E [similarity] - Methan

C:Species: Methanococcus jannaschii

C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 02-Mar-2001

C:Accession: G64405

R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,

J.; Reisch, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;

rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.

Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Bordovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C

A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii

A:Reference number: A64300; MUID:96337999; PMID:8686087

A:Accession: G64405
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-303 <BUL>
A:Cross-references: GB:067529; GB:L77117; NID:g1591532; PIDN:AAB98652.1; PID:g1591534
C:Genetics:
A:Map position: FOR773706-774617
C:Superfamily: Methanobacterium thermoautotrophicum tetrahydromethanopterin S-methylt
C:Keywords: methyltransferase

Query Match 33.3%; Score 5; DB 2; Length 303;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SHLGP 5
Db 131 SHLGP 135

RESULT 148

B70835

hypothetical protein Rv0276 - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999

C:Accession: B70835

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon

Rajandream, M.A.; Rogers, J.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd,

Nature 393, 537-544, 1998

A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: B70835

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-306 <COL>

A:Cross-references: GB:AL021930; GB:AL123456; NID:g3261524; PIDN:CAA17351.1; PID:e125

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: Rv0276

Query Match 33.3%; Score 5; DB 2; Length 306;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 8 STPEs 12
Db 120 STPEs 124

RESULT 149

H64819

formate acetyltransferase activating enzyme (EC 1.97.1.4) 3 - Escherichia coli (stral

C:Species: Escherichia coli

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 01-Mar-2002

C:Accession: H64819

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;

A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: H64819

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-308 <BLAT>

A:Cross-references: GB:AE000184; GB:U00096; NID:g1787036; PIDN:AAC73911.1; PID:g17870

A:Experimental source: strain K-12, substrain MG1655

C:Genetics:

A:Gene: ybiY

C:Function:

A:Description: activation of pyruvate formate-lyase under anaerobic conditions by gen

A:Pathway: anaerobic glucose metabolism

A:Note: Iron dependent

Query Match 33.3%; Score 5; DB 2; Length 299;
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRA 14
 |||||
 DB 35 PESRA 39

RESULT 141

transcription factor Brn-5 - human
 N:Alternate names: homeobox protein mPOU; TCR beta enhancer binding protein TCFbeta1
 C:Species: Homo sapiens (man)
 C>Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 22-Jun-1999
 C:Accession: A54687; S42575; S40151
 R:Reisler, H.; Brickner, H.; Galikwad, J.; Fotedar, A.
 Mol. Cell. Biol. 13, 5450-5460, 1993
 A>Title: A novel POU domain protein which binds to the T-cell receptor beta enhancer.
 A:Reference number: A54687; MUID:93360980; PMID:8102789
 A:Accession: A54687
 A:Molecule type: mRNA
 A:Residues: 1-301 <MES>
 A:Cross-references: GB:L14482
 R:Wey, E.; Lyons, G.E.; Schaefer, B.W.
 Eur. J. Biochem. 220, 753-762, 1994
 A>Title: A human POU domain gene, mPOU, is expressed in developing brain and specific ad
 A:Reference number: S42575; MUID:94192665; PMID:7908264
 A:Accession: S42575
 A:Molecule type: mRNA
 A:Residues: 1-301 <MEY>
 A:Cross-references: EMBL:Z21966; NID:g437806; PIDN:CAA79977.1; PID:g437807
 C:Genetics:
 A:Gene: GDB:POU6F1; BRN5; MPOU; TCFB1
 A:Cross-references: GDB:361078
 A:Map position: 12pter-12qter
 C:Superfamily: transcription factor Brn-5; homeobox homology; POU domain homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F:146-213/Domain: POU domain homology <POU>
 F:235-291/Domain: homeobox homology <HOX>

Query Match 33.3%; Score 5; DB 1; Length 301;
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STPES 12
 |||||
 DB 74 STPES 78

RESULT 142

transcription factor Brn-5 - rat
 N:Alternate names: homeobox protein mPOU; TCR beta enhancer binding protein TCFbeta1
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 24-Oct-1997
 C:Accession: A48880
 R:Andersen, B.; Schonemann, M.D.; Pearce II, R.V.; Jenne, K.; Sugarman, J.; Rosenfeld, M.
 J. Biol. Chem. 268, 23390-23398, 1993
 A>Title: Brn-5 is a divergent POU domain factor highly expressed in layer IV of the neo
 A:Reference number: A48880; MUID:94043133; PMID:7901208
 A:Accession: A48880
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-301 <AND>
 A:Cross-references: GB:I23204; NID:g349723
 A:Experimental source: anterior pituitary
 A>Note: sequence extracted from NCBI backbone (NCBIN:138920, NCBIPI:138921)
 C:Superfamily: transcription factor Brn-5; homeobox homology; POU domain homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F:146-213/Domain: POU domain homology <POU>
 F:235-291/Domain: homeobox homology <HOX>

Query Match 33.3%; Score 5; DB 1; Length 301;
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STPES 12
 |||||
 DB 74 STPES 78

RESULT 143

8-oxoguanine DNA glycosylase [imported] - Halobacterium sp. NRC-1
 B4282
 C:Species: Halobacterium sp. NRC-1
 C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 01-Mar-2002
 C:Accession: B4282
 R:Ng, W.Y.; Kennedy, S.P.; Mahairas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky
 ; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Ja
 Jung, K.H.; Alam, M.; Freitas, T.
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
 A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.;
 A>Title: Genome sequence of Halobacterium species NRC-1
 A:Reference number: B4282; MUID:20504483; PMID:11016950
 A:Accession: B4282
 A:Molecule type: DNA
 A>Status: preliminary
 A:Residues: 1-301 <STO>
 A:Cross-references: GB:AE004437; NID:g10580792; PIDN:AMG19622.1; GSPDB:GN00138
 C:Genetics:
 A:Gene: ogg
 C:Superfamily: mouse 8-oxoguanine DNA-glycosylase

Query Match 33.3%; Score 5; DB 2; Length 301;
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HLGPH 6
 |||||
 DB 274 HLGPH 278

RESULT 144

hypothetical protein T19F6_80 - Arabidopsis thaliana
 T13457
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 04-Mar-2000
 C:Accession: T13457
 R:Beyan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.;
 submitted to the Protein Sequence Database, July 1999
 A:Reference number: T13457
 A:Accession: T13457
 A:Molecule type: DNA
 A:Residues: 1-302 <BEV>
 A:Cross-references: EMBL:AL109619; GSPDB:GN00062; ATSP:T19F6_80
 A:Experimental source: cultivar Columbia; BAC clone T19F6
 C:Genetics:
 A:Gene: ATSP:T19F6_80
 A:Map position: 4
 A:Introns: 129/3; 244/1
 C:Superfamily: Arabidopsis thaliana hypothetical protein T19F6_80

Query Match 33.3%; Score 5; DB 2; Length 302;
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRAA 15
 |||||
 DB 124 ESRAA 128

RESULT 145

hypothetical protein Vng1779c [imported] - Halobacterium sp. NRC-1
 F4329

```
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 11 ESRRA 15
    |||||
Db 166 ESRRA 170

RESULT 136
AF0098
probable exported protein YPO0800 [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C:Accession: AF0098
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
ll, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AF0098
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-295 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC89649.1; PID:915978877; GSPDB:GN00175
C:Genetics:
A:Gene: YPO0800

Query Match 33.3%; Score 5; DB 2; Length 295;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 9 TPESR 13
    |||||
Db 202 TPESR 206

RESULT 137
T13317
hypothetical protein 28 - Streptococcus phage phi-O1205
C:Species: Streptococcus phage phi-O1205
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 24-Sep-1999
C:Accession: T13317
R:Stanley, E.; Fitzgerald, G.F.; Le Marrec, C.; Fayard, B.; van Sinderen, D.
Microbiology 143, 3417-3429, 1997
A:Title: Sequence analysis and characterization of phi O1205, a temperate bacteriophage
A:Reference number: Z17654; MUID:98048466; PMID:9387220
A:Accession: T13317
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-297 <STA>
A:Cross-references: EMBL:U88974; NID:Q2444080; PID:Q2444107; PIDN:AAC79543.1
A:Experimental source: host Streptococcus thermophilus strain CNR21205

Query Match 33.3%; Score 5; DB 2; Length 297;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 10 PESRA 14
    |||||
Db 251 PESRA 255

RESULT 138
F90741
hypothetical protein ECS0902 [imported] - Escherichia coli (strain O157:H7, substrain RI
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: F90741
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gawawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hatori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gend
```

```
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: F90741
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-299 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA834325.1; PID:913360361; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: ECS0902
C:Superfamily: probable pyruvate formate-lyase 2 activating enzyme; ferredoxin 2[4fe-

Query Match 33.3%; Score 5; DB 2; Length 299;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 10 PESRA 14
    |||||
Db 35 PESRA 39

RESULT 139
AB5592
hypothetical protein yblY [imported] - Escherichia coli (strain O157:H7, substrain ED
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: AB5592
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Diallanta, E.; Potamousis, K.; Apoda
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: AB5480; MUID:21074935; PMID:11206551
A:Accession: AB5592
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-299 <STO>
A:Cross-references: GB:AE005174; NID:q12513827; PIDN:AAG55197.1; GSPDB:GN00145; UWGP:
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: yblY
C:Superfamily: probable pyruvate formate-lyase 2 activating enzyme; ferredoxin 2[4fe-

Query Match 33.3%; Score 5; DB 2; Length 299;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 10 PESRA 14
    |||||
Db 35 PESRA 39

RESULT 140
AH0602
probable formate acetyltransferase activating enzyme (EC 1.97.1.4) [imported] - Salmo
C:Species: Salmonella enterica subsp. enterica serovar Typhl
A:Note: This species has also been called Salmonella typhl
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 03-Jun-2002
C:Accession: AH0602
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
th, T.; Connerton, P.; Cronin, A.; Davies, P.; Davies, R.M.; Dowd, L.; White, N.; Farr
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se
A:Reference number: AB0502; PMID:11677608
A:Accession: AH0602
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-299 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD05290.1; PID:916502054; GSPDB:GN00176
C:Genetics:
A:Gene: yblY
C:Superfamily: probable pyruvate formate-lyase 2 activating enzyme; ferredoxin 2[4fe-
C:Keywords: oxidoreductase
```

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
J.
Nature 402, 761-768, 1999
A>Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
A:Reference number: AB4420; MUID:20083487; PMID:10617197
A:Accession: B84807
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1286 <STO>
A:Cross-references: GB:AE002093; NID:g3786011; PIDN:AA67357.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g38610
A:Map position: 2

Query Match 33.3%; Score 5; DB 2; Length 286;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
|||||
DB 22 RSTPE 26

RESULT 132
T15779
hypothetical protein C35B8.1 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 24-Nov-1999
C:Accession: T15779
R;Leimbach, D.
Submitted to the EMBL Data Library, March 1995
A:Description: The sequence of *C. elegans* cosmid C35B8.
A:Reference number: Z18401
A:Accession: T15779
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-287 <LEI>
A:Cross-references: EMBL:U03520; NID:g746532; PID:g746533; PIDN:AA64653.1; CESP:C35B8.1
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:C35B8.1
A:Introns: 45/3
C:Superfamily: unassigned collagens

Query Match 33.3%; Score 5; DB 2; Length 287;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRAA 15
|||||
DB 93 ESRAA 97

RESULT 133
G70605
probable hydrolase - *Mycobacterium tuberculosis* (strain H37Rv)
C:Species: *Mycobacterium tuberculosis*
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: G70605
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,
S.; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.;
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A>Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: G70605
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-291 <COL>
A:Cross-references: GB:Z92774; GB:AL123456; NID:g3261729; PIDN:CA807143.1; PID:g1877300
A:Experimental source: strain H37Rv

C:Genetics:
A:Gene: RV3569c
C:Superfamily: tropinesterase

Query Match 33.3%; Score 5; DB 2; Length 291;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPE 12
|||||
DB 197 STPE 201

RESULT 134
T00996
En/Spm-like transposon protein [imported] - *Arabidopsis thaliana*
N:Alternate names: hypothetical protein T9J22.30
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 23-Mar-2001
C:Accession: T00996; H84662
R;Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; K
Submitted to the EMBL Data Library, April 1998
A:Description: *Arabidopsis thaliana* chromosome II BAC T9J22 genomic sequence.
A:Reference number: Z14161
A:Accession: T00996
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-292 <ROU>
A:Cross-references: EMBL:AC002505; NID:g2739359; PID:g2739387
A:Experimental source: cultivar Columbia
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
J.
Nature 402, 761-768, 1999
A>Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
A:Reference number: AB4420; MUID:20083487; PMID:10617197
A:Accession: H84662
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-292 <STO>
A:Cross-references: GB:AE002093; NID:g2739387; PIDN:AA614510.1; GSPDB:GN00139
C:Genetics:
A:Gene: T9J22.30; At2g26630
A:Map position: 2
A:Introns: 43/1

Query Match 33.3%; Score 5; DB 2; Length 292;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPFR 7
|||||
DB 135 LGPFR 139

RESULT 135
AA3663
host-inducible protein A - *Rhizobium fredii*
C:Species: *Rhizobium fredii*
C>Date: 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change 04-Mar-2000
C:Accession: AA3663
R;Sadovsky, M.J.; Olson, E.R.; Foster, V.E.; Kossiak, R.M.; Verma, D.P.S.
J. Bacteriol. 170, 1771-178, 1988
A>Title: Two host-inducible genes of *Rhizobium fredii* and characterization of the *inc*
A:Reference number: AA3663; MUID:88086864; PMID:2447061
A:Accession: AA3663
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-295 <SAD>
A:Cross-references: GB:M19019; NID:g152247; PIDN:AA26294.1; PID:g152248
C:Superfamily: *Rhizobium fredii* host-inducible protein A
Query Match 33.3%; Score 5; DB 2; Length 295;

J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of *Methanobacterium thermoautotrophicum* Delta H: funct
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: B69025
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-278 <MTH>
A:Cross-references: GB:AE000887; GB:AE000666; NID:g2622289; PIDN:AAB85675.1; PID:g262229
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH186
C:Superfamily: conserved hypothetical protein MJ1157

Query Match 33.3%; Score 5; DB 1; Length 278;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STRPS 12
DB 184 STRPS 188

RESULT 127
E83456
hypothetical protein PA1522 [imported] - *Pseudomonas aeruginosa* (strain PA01)
C:Species: *Pseudomonas aeruginosa*
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: E83456
R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Laidig, K.; Lam,
..Loiy, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: E83456
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-279 <STO>
A:Cross-references: GB:AE004580; GB:AE004091; NID:g9947468; PIDN:MAG04911.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA1522

Query Match 33.3%; Score 5; DB 2; Length 279;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESR 13
DB 270 TPESR 274

RESULT 128
C70642
Probable ribosomal protein L2 rplB - *Mycobacterium tuberculosis* (strain H37RV)
C:Species: *Mycobacterium tuberculosis*
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: C70642
R:Conor, S.T.; Davies, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Rajandream, M.A.; Rogers, R.; Devlin, K.; Felkell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Nature 393, 537-544, 1998
A:Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: C70642
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-280 <COL>
A:Cross-references: GB:Z84395; GB:AL123456; NID:g3261698; PIDN:CAB06467.1; PID:g1806172
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: rplB

C:Superfamily: *Escherichia coli* ribosomal protein L2

Query Match 33.3%; Score 5; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
DB 26 RSTPE 30

RESULT 129
T28857
hypothetical protein R03E9.1 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 18-Feb-2000
C:Accession: T28857
R:Wilcox, L.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid R03E9.
A:Reference number: Z20533
A:Accession: T28857
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-281 <WIL>
A:Cross-references: EMBL:U40947; PIDN:AAC48067.1; GSPDB:GN00028; CESP:R03E9.1
A:Experimental source: strain Bristol N2; clone R03E9
C:Genetics:
A:Gene: CESP:R03E9.1
A:Map position: X
A:introns: 52/1; 145/3

Query Match 33.3%; Score 5; DB 2; Length 281;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRA 14
DB 197 PESRA 201

RESULT 130
T13621
hypothetical protein gp284 - *Streptococcus phage phi-sfil1*
C:Species: *Streptococcus phage phi-sfil1*
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 13-Aug-1999
C:Accession: T13621
R:Lucchini, S.; Desiere, F.; Brunsow, H.
Virology 246, 63-73, 1998
A:Title: The structural gene module in *Streptococcus thermophilus* bacteriophage phi s
A:Reference number: Z17696; MUID:98321150; PMID:9656994
A:Accession: T13621
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-284 <LUC>
A:Cross-references: EMBL:AF057033; NID:g3320432; PID:g3320434; PIDN:AAC34398.1
A:Experimental source: specific host *Streptococcus thermophilus*

Query Match 33.3%; Score 5; DB 2; Length 284;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRA 14
DB 238 PESRA 242

RESULT 131
B84807
probable RNA-binding protein [imported] - *Arabidopsis thaliana*
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: B84807

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 8 STPES 12
 |||||
 Db 166 STPES 170

RESULT 122

H97543
 hypothetical protein AGR_C_2792 [imported] - Agrobacterium tumefaciens (strain C58, Cere
 C:Species: Agrobacterium tumefaciens
 C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
 C:Accession: H97543
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Oucollo, B.; Goldman,
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
 Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
 A:Reference number: A97359; PMID:11743194
 A:Accession: H97543
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-271 <KUR>
 A:Cross-references: GB:AE007869; PIDN:AAK87305.1; PID:g15156600; GSPDB:GN00169
 C:Genetics:
 A:Gene: AGR_C_2792
 A:Map position: circular chromosome

Query Match 33.3%; Score 5; DB 2; Length 271;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGP 5
 |||||
 Db 111 SHLGP 115

RESULT 123

AG2215
 hypothetical protein all3278 [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp.
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
 C:Accession: AG2215
 R:Kameo, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi,
 Nazareki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AG2215
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-272 <KUR>
 A:Cross-references: GB:BA000019; PIDN:BAW4977.1; PID:g17132373; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: all3278

Query Match 33.3%; Score 5; DB 2; Length 272;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 TPESR 13
 |||||
 Db 111 TPESR 115

RESULT 124

T22993
 hypothetical protein F59C6.8 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T22993
 R:Wilkinson, J.

submitted to the EMBL Data Library, August 1996
 A:Reference number: Z19648

A:Accession: T22993
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-274 <NLL>
 A:Cross-references: EMBL:Z79600; PIDN:CAB01879.1; GSPDB:GN00019; CESP:F59C6.8
 A:Experimental source: clone F59C6
 C:Genetics:
 A:Gene: CESP:F59C6.8
 A:Map position: 1
 A:Introns: 51/3; 85/3; 97/3; 170/3; 193/1; 211/3

Query Match 33.3%; Score 5; DB 2; Length 274;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRAA 15
 |||||
 Db 110 ESRAA 114

RESULT 125

G69963
 lipoprotein SpoIIIV-like homolog yqjG - Bacillus subtilis
 C:Species: Bacillus subtilis
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
 C:Accession: G69963
 R:Kunst, F.; Ogatawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Be
 C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
 A.; Ehlich, S.D.; Emmerison, P.T.; Entian, K.D.; Erlington, J.; Fabbet, C.; Ferrari,
 Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gal
 lech, J.; Harwood, C.R.; Henault, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, I
 Koelter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinc
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Ma
 Y., M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portier
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scan
 A:Authors: Schleich, S.; Schreier, R.; Scoffone, F.; Sekiguchi, J.; Sekoska, A.; Si
 akench, M.; Tamakoshi, A.; Tanaka, T.; Tempstra, P.; Tognoni, A.; Tosato, V.; Uchly
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshid
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtil
 A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: G69963
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-275 <RUN>
 A:Cross-references: GB:Z99116; GB:AL009126; MID:g2634723; PIDN:CAB14320.1; PID:g2634
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: yqjG
 C:Superfamily: stage III sporulation protein; stage III sporulation protein homology
 F:43-257/Domain: stage III sporulation protein homology <SPOR>

Query Match 33.3%; Score 5; DB 1; Length 275;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPE 11
 |||||
 Db 160 RSTPE 164

RESULT 126

B69025
 conserved hypothetical protein MTH1186 - Methanobacterium thermoautotrophicum (strai
 C:Species: Methanobacterium thermoautotrophicum
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
 C:Accession: B69025
 R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, J
 ; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani,
 Kl, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

C:Genetics:
A:Gene: Atu1514
A:Map position: circular chromosome

Query Match 33.3%; Score 5; DB 2; Length 266;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHGP 5
|||||
DB 106 SHGP 110

RESULT 118

E97685
monofunctional biosynthesis peptidoglycan transglycosylase-like protein (AF287157) [ImpC
C:Species: Agrobacterium tumefaciens
C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C:Accession: E97685
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
A>Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; PMID:11743194
A:Accession: E97685
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-267 <KUR>
A:Cross-references: GB:AE007869; PIDN:AAK88438.1; PID:g15157935; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C_4930
A:Map position: circular chromosome

Query Match 33.3%; Score 5; DB 2; Length 267;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPE 12
|||||
DB 22 STPE 26

RESULT 119

AD0898
PIS-transport family phosphotransfer protein STY3438 [Imported] - Salmonella enterica
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A>Note: this species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
C:Accession: AD0898
R:Parthill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, Th. T.; Connerton, P.; Cronin, A.; Davies, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Mole, S.; O'Garra, P.; Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A>Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; PMID:11677608
A:Accession: AD0898
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-268 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD07778.1; PID:g16504327; GSPDB:GN00176
C:Genetics:
A:Gene: STY3438

Query Match 33.3%; Score 5; DB 2; Length 268;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESR 13
|||||
DB 63 TPESR 67

RESULT 120

G75411
probable transposase - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 26-May-2000
C:Accession: G75411; G75638
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.; Science 286, 1571-1577, 1999
A>Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: G75411
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-270 <WH1>
A:Cross-references: GB:AE001977; GB:AE000513; NID:g6459045; PIDN:AAF10868.1; PID:g645
A:Experimental source: strain R1
A:Accession: G75638
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-270 <WH2>
A:Cross-references: GB:AE001827; NID:g6460959; PIDN:AAF2671.1; PID:g6460968; TIGR:DR
A:Experimental source: strain R1
C:Genetics: <WH1A>
A:Gene: DR1296
A:Map position: 1
C:Genetics: <WH2B>
A:Gene: DR0033
A:Genome: plasmid
A>Note: plasmid Cpl
C:Superfamily: Synechocystis transposase s111710

Query Match 33.3%; Score 5; DB 2; Length 270;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPE 12
|||||
DB 114 STPE 118

RESULT 121

C96010
conserved hypothetical protein SMB21517 [Imported] - Sinorhizobium meliloti (strain 1
C:Species: Sinorhizobium meliloti
C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: C96010
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Her Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A>Title: The complete sequence of the 1,683-kb psymb megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: C96010
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-271 <KUR>
A:Cross-references: GB:AL591985; PIDN:CAC49747.1; PID:g15141234; GSPDB:GN00167
A:Experimental source: strain 1021, megaplasmid psymb
R:Galibert, F.; Finan, T.M.; Long, S.R.; Punler, A.; Abola, P.; Ampe, F.; Barloy-Hubl
Pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R. L.; Hyman, R.W.; Jones, T.; Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lejau
hebaul, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, A>Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: SMB21517
A:Genome: plasmid

Query Match 33.3%; Score 5; DB 2; Length 271;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;

Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 7 RSTPE 11
Db 134 RSTPE 138

RESULT 113

T16924
hypothetical protein T22E5.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 04-Mar-2000
C:Accession: T16924
R:Minx, P.
Submitted to the EMBL Data Library, December 1995
A:Description: The sequence of C. elegans cosmid T22E5.
A:Reference number: Z18606
A:Accession: T16924
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-249 <MIN>
A:Cross-references: EMBL:U03282; NID:g1125825; PID:g1125830; PIDN:AAA83617.1; CESP:T22E5
C:Genetics:
A:Gene: CESP:T22E5.6
A:Introns: 53/3; 99/1; 149/1; 177/3
C:Superfamily: Caenorhabditis elegans hypothetical protein T22E5.6

Query Match 33.3%; Score 5; DB 2; Length 249;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPE 12
Db 7 STPE 11

RESULT 114

B48725
MDV specific protein - Marek's disease virus
C:Species: Marek's disease virus
C:Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 23-Mar-1995
C:Accession: B48725
R:Camp, H.S.; Silva, R.F.; Coussens, P.M.
Virology 196, 484-495, 1993
A:Title: Defective Marek's disease virus DNA contains a gene encoding a potential nucle
A:Reference number: A48725; MUID:9383373; PMID:8396799
A:Accession: B48725
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-252 <CAM>
A:Experimental source: serotype 2, 281M1/1
A:Note: sequence inconsistent with nucleotide translation
A:Note: sequence extracted from NCBI backbone (NCBIN:137884, NCBIP:137886)

Query Match 33.3%; Score 5; DB 2; Length 252;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRA 14
Db 64 PESRA 68

RESULT 115

AG2910
hypothetical protein mtga [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: AG2910
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavln, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kau
ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AG2910
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-263 <KUR>
A:Cross-references: GB:AE008688; PIDN:AA43701.1; PID:g17741229; GSPDB:GN00186
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: mtga
A:Map position: circular chromosome

Query Match 33.3%; Score 5; DB 2; Length 263;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPE 12
Db 18 STPE 22

RESULT 116

AB7345
transcription regulator, TetR family [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: AB7345
R:Nieman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,
B.; Laub, M.T.; Debby, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; K
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: AB7249; MUID:21173698; PMID:11259647
A:Accession: AB7345
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-265 <STO>
A:Cross-references: GB:AE005673; NID:g13422011; PIDN:AAK22757.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC0772

Query Match 33.3%; Score 5; DB 2; Length 265;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
Db 188 RSTPE 192

RESULT 117

A12762
conserved hypothetical protein Atul1514 [imported] - Agrobacterium tumefaciens (strain
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: A12762
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Wo
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavln, T.; Levy, R.; Li, M.; McCle
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kau
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: A12762
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-266 <KUR>
A:Cross-references: GB:AE008688; PIDN:AA42519.1; PID:g17739393; GSPDB:GN00186
A:Experimental source: strain C58 (Dupont)

RESULT 108
H69030
coenzyme PQQ synthesis protein III - Methanobacterium thermoautotrophicum (strain Delta
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: H69030
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
; Qiu, D.; Spadofora, R.; Vitale, R.; Wang, Y.; Mierowski, J.; Gibson, R.; Jivani, N.
K, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func
A:Reference number: A69000; MUID:96037514; PMID:9371463
A:Accession: H69030
A:Molecule type: DNA
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Cross-references: GB:AE000890; GB:AE000666; NID:g2622331; PIDN:AA85716.1; PID:g262233
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH1227
A:Start codon: GTG
A:Keywords: iron; metalloprotein
F:34,38,41/Binding site: iron (Cys) #status predicted
Query Match 33.3%; Score 5; DB 2; Length 247;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 9 TPESR 13
|||||
DB 43 TPESR 47

RESULT 109
T17311
hypothetical protein DKFZP434C128.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T17311
R:Ottensmider, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, September 1999
A:Reference number: Z18726
A:Accession: T17311
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-247 <OFT>
A:Cross-references: EMBL:AL117578
A:Experimental source: adult testis; clone DKFZP434C128
C:Genetics:
A:Note: DKFZP434C128.1
Query Match 33.3%; Score 5; DB 2; Length 247;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 7 RSTPE 11
|||||
DB 112 RSTPE 116

RESULT 110
B97669
hypothetical protein AGR_C_4681 [Imported] - Agrobacterium tumefaciens (strain C58, Cere
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C:Accession: B97669
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Mollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; PMID:11743194
A:Accession: B97669
A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-247 <KUR>
A:Cross-references: GB:AE007869; PIDN:AAK8307.1; PID:g15157779; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C_4681
A:Map position: circular chromosome
Query Match 33.3%; Score 5; DB 2; Length 247;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 3 LGPHR 7
|||||
DB 95 LGPHR 99

RESULT 111
AG2893
conserved hypothetical protein Atu2584 [Imported] - Agrobacterium tumefaciens (strain
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: AG2893
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutlyavin, T.; Levy, R.; Li, M.; McCl
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AG2893
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-247 <KUR>
A:Cross-references: GB:AE008688; PIDN:ALA43565.1; PID:g17741079; GSPDB:GN00186
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu2584
A:Map position: circular chromosome
Query Match 33.3%; Score 5; DB 2; Length 247;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 3 LGPHR 7
|||||
DB 95 LGPHR 99

RESULT 112
A70745
probable transcription regulator Ry0494 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: A70745
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: A70745
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-249 <COL>
A:Cross-references: GB:Z77162; GB:AL134346; NID:g3261606; PIDN:CAB00955.1; PID:g14492
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Ry0494
C:Superfamily: regulatory protein uxur 2
Query Match 33.3%; Score 5; DB 2; Length 249;

sporulated oocyst antigen T4 precursor - *Eimeria tenella* (fragment)
C:Species: *Eimeria tenella*
C>Date: 09-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 07-May-1999
C:Accession: A54501
R:Brothers, V.M.; Kuhn, I.; Paul, L.S.; Gabe, J.D.; Andrews, W.H.; Stas, S.R.; McCaman, Moll. Biochem. Parasitol. 28, 235-248, 1988
A:Title: Characterization of a surface antigen of *Eimeria tenella* sporozoites and synthesis of monoclonal antibodies
A:Reference number: A54501; PMID:88261435; PMID:3290678
A:Accession: A54501
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-230 <PRO>
A:Cross-references: GB:M21004
C:Keywords: disulfide bond

Query Match 33.3%; Score 5; DB 2; Length 230;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRA 15
Db 44 ESRA 48

RESULT 104
C72576
Probable glutamine transport ATP-binding protein APE1891 - *Aeropyrum pernix* (strain K1)
C:Species: *Aeropyrum pernix*
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 08-Sep-2000
C:Accession: C72576
R:Kawababayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; K DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, *Aeropyrum pernix* strain K1
A:Reference number: A72450; PMID:99310359; PMID:10382966
A:Accession: C72576
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-238 <KAW>
A:Cross-references: DDBJ:AP000062; NID:95105244; PIDN:BA80896.1; PID:95105583
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE1891
C:Superfamily: short-chain ATP-binding cassette proteins; ATP-binding cassette homology F.10-206/Domain: ATP-binding cassette homology <ABC>

Query Match 33.3%; Score 5; DB 2; Length 238;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRA 14
Db 163 PESRA 167

RESULT 105
B88115
protein F53C3.4 [imported] - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: B88115
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biology and the genetics of a model organism
A:Reference number: A75000; PMID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_elegans/ for the complete genome sequence
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and Science 283, 2103, 1999
A:Accession: B88115
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-244 <STO>
A:Cross-references: GB:chr_II; PIDN:AC67455.1; PID:g3786482; GSPDB:GN00020; CESP:F53C3.4
C:Genetics:

A:Gene: F53C3.4
A:Map position: 2

Query Match 33.3%; Score 5; DB 2; Length 244;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPE 11
Db 204 RSTPE 208

RESULT 106
AG2071
hypothetical protein alr2125 [imported] - *Nostoc* sp. (strain PCC 7120)
C:Species: *Nostoc* sp.
A:Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AG2071
R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kurita, T.; Sasamoto, S.; Watanabe, A.; Irigun, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Masuda, M.; Tabata DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Nostoc* sp. strain PCC 7120
A:Reference number: AB1807; PMID:21595285; PMID:11759840
A:Accession: AG2071
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-244 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA873824.1; PID:g17131216; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr2125

Query Match 33.3%; Score 5; DB 2; Length 244;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STPE 12
Db 208 STPE 212

RESULT 107
T33469
hypothetical protein F43B10.1 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 04-Mar-2000
C:Accession: T33469
R:Fulton, R.; Hawkins, J.; Rohlfing, T.
Submitted to the EMBL Data Library, October 1998
A:Description: The sequence of *C. elegans* cosmid F43B10.
A:Reference number: Z21351
A:Accession: T33469
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-247 <FUL>
A:Cross-references: EMBL:AF098500; PIDN:AC67400.1; GSPDB:GN00028; CESP:F43B10.1
A:Map position: X
A:Introns: 86/2; 103/3; 128/3
C:Superfamily: *Caenorhabditis elegans* hypothetical protein F43B10.1

Query Match 33.3%; Score 5; DB 2; Length 247;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGP 5
Db 162 SHLGP 166

A:Reference number: A86909; MUID:21128732; PMID:11234002
A:Accession: A87094
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-224 <SNO>
A:CROSS-references: GB:AL450380; NID:q13093330; PIDN:CAC30430.1; GSPDB:GN00147
C:Genetics:
A:Gene: c1p
C:Superfamily: endopeptidase Clp chain p

Query Match 33.3%; Score 5; DB 2; Length 224;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESR 13
DB 17 TPESR 21

RESULT 99
D85018
probable hypersensitive response protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C:Accession: D85018
R:Anonymous, The European Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617158
A:Accession: D85018
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-227 <STO>
A:CROSS-references: GB:NC_001268; NID:g7267638; PIDN:CAB80950.1; GSPDB:GN00140
C:Genetics:
A:Gene: AT4g01410
A:Map position: 4

Query Match 33.3%; Score 5; DB 2; Length 227;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPES 12
DB 24 STPES 28

RESULT 100
D95865
probable pentose-5-phosphate-3-epimerase protein (EC 5.1.3.-) [imported] - Sinorhizobium
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: D95865
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1,663-kb pSymb megaplasmid from the N2-fixing endo
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: D95865
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-228 <KUR>
A:CROSS-references: GB:AL591985; PIDN:CAC48588.1; PID:q15140060; GSPDB:GN00167
A:Experimental source: strain 1021, megaplasmid pSymb
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kaiman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
Hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:

A:Gene: ppe; SMD20195
A:Genome: plasmid
C:Superfamily: Yeast ribulose-5-phosphate-epimerase
C:Keywords: isomerase

Query Match 33.3%; Score 5; DB 2; Length 228;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPES 12
DB 120 STPES 124

RESULT 101
S41043
rum1 protein - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 31-Mar-1992 #sequence_revision 14-Sep-1994 #text_change 29-Oct-1999
C:Accession: S41043
R:Moreno, S.; Nurse, P.
Nature 367, 236-242, 1994
A:Title: Regulation of progression through the G1 phase of the cell cycle by the rum1
A:Reference number: S41043; MUID:94166876; PMID:8121488
A:Accession: S41043
A:Molecule type: DNA
A:Residues: 1-230 <MOR>
A:CROSS-references: EMBL:X77730; NID:g456668; PIDN:CAA54786.1; PID:g456669
C:Genetics:
A:Gene: rum1
A:Map position: 2
C:Superfamily: Schizosaccharomyces rum1 protein

Query Match 33.3%; Score 5; DB 2; Length 230;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPES 12
DB 15 STPES 19

RESULT 102
T40233
Rum1p - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
C:Accession: T40233
R:Moreno, S.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, June 1998
A:Reference number: 221915
A:Accession: T40233
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-230 <MOR>
A:CROSS-references: EMBL:AL023796; PIDN:CAA19370.1; GSPDB:GN00067; SPDB:SPBC32F12.09
A:Experimental source: strain 972h-; cosmid c32F12
C:Genetics:
A:Gene: SPDB:SPBC32F12.09
A:Map position: 2
C:Superfamily: Schizosaccharomyces rum1 protein

Query Match 33.3%; Score 5; DB 2; Length 230;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPES 12
DB 15 STPES 19

RESULT 103
A54501

RESULT 94
F82627
phosphoglycerate mutase XP1886 [imported] - *Xylella fastidiosa* (strain 9a5c)
C:Species: *Xylella fastidiosa*
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 17-Nov-2000
C:Accession: F82627
R:Anonymous, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: For a complete list of authors see reference number A59328 below
A:Accession: F82627
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-214 <SIM>
A:Cross-references: GB:AE004008; GB:AE003849; NID:g9106961; PIDN:AAF84692.1; GSPDB:GN001
A:Experimental source: strain 9a5c
R:Simpton, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Brienes, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carre, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, B
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmeri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuchiko, M.H.; Valada, H.; Van Sluys, W.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XP1886
C:Superfamily: Aquifex aeolicus phosphoglycerate mutase; phosphoglycerate mutase homolog

Query Match 33.3%; Score 5; DB 2; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 11 ESRRA 15
DB 68 ESRRA 72
RESULT 95
T16393
hypothetical protein F48B9.5 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 02-Sep-2000
C:Accession: T16393
R:Miller, N.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid F48B9.
A:Reference number: Z18507
A:Accession: T16393
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-215 <MIT>
A:Cross-references: EMBL:U40955; NID:g1072262; PID:g1072263; PIDN:AAA81747.1; CESP:F48B9
C:Genetics:
A:Gene: CESP:F48B9.5
A:Introns: 23/3; 100/1; 127/3; 152/2; 181/2
C:Superfamily: paired box homology
F:97215/Domain: paired box homology <PBH>

Query Match 33.3%; Score 5; DB 2; Length 215;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 8 STPES 12
DB 9 STPES 13

RESULT 96
HB3475
Probable transcription regulator PA1359 [imported] - *Pseudomonas aeruginosa* (strain P
C:Species: *Pseudomonas aeruginosa*
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: HB3475
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;
Adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; L
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pa
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: HB3475
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-218 <STO>
A:Cross-references: GB:AE004565; GB:AE004091; NID:g9947294; PIDN:AAG04748.1; GSPDB:GN
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA1359

Query Match 33.3%; Score 5; DB 2; Length 218;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 10 PESRA 14
DB 124 PESRA 128

RESULT 97
T42605
envelope protein - equine herpesvirus 4 (strain NS80567)
C:Species: equine herpesvirus 4
A:Variety: strain NS80567
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
C:Accession: T42605
R:Reford, E.A.; Watson, M.S.; Perry, J.; Cullinane, A.A.; Davison, A.J.
J. Gen. Virol. 79, 1197-1203, 1998
A:Title: The DNA sequence of equine herpesvirus-4.
A:Reference number: Z2173; MUID:98264497; PMID:9603335
A:Accession: T42605
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-219 <TEL>
A:Cross-references: EMBL:AF030027; NID:g2605950; PIDN:AAC59581.1; PID:g2606009
A:Experimental source: strain NS80567
C:Genetics:
A:Gene: 62
C:Superfamily: varicella-zoster virus gene 60 protein

Query Match 33.3%; Score 5; DB 2; Length 219;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 8 STPES 12
DB 190 STPES 194

RESULT 98
A87094
ATP-dependent Clp protease proteolytic subunit [imported] - *Mycobacterium leprae*
C:Species: *Mycobacterium leprae*
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 27-Nov-2001
C:Accession: A87094
R:Coile, S.T.; Eigmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.;
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.;
A:Title: Massive gene decay in the leprosy bacillus.

probable heme transport protein CCB206 - rape mitochondrion
 M:Alternate names: HelB-homolog
 C:Species: mitochondrion Brassica napus (rape)
 C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 18-Aug-2000
 C:Accession: T09503
 R:Itami, K.; Handa, H.
 Curr. Genet. 34, 318-325, 1998
 A:Title: Rapeseed mitochondrial ccb206, a gene involved in cytochrome c biogenesis, is c
 206 locus.
 A:Reference number: 216701; MUID:99015983; PMID:9799366
 A:Accession: T09503
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-206 <HAN>
 A:Cross-references: EMBL:DJ3697
 C:Genetics:
 A:Gene: ccb206
 A:Genome: mitochondrion
 C:Function:
 A:Description: Involved in cytochrome c biogenesis
 C:Superfamily: cytochrome c biogenesis protein CycW
 C:Keywords: mitochondrion; RNA editing

Query Match 33.3%; Score 5; DB 2; Length 206;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGP 5
 DB 46 SHLGP 50

RESULT 90

AF0498
 hypothetical protein YPO4106 [imported] - Yersinia pestis (strain CO92)
 C:Species: Yersinia pestis
 C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
 C:Accession: AF0498
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tlball, R.W.; Holden, M.T.G.; Prentice, M.B.
 deno-Tarrage, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
 11. M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett,
 Nature 413, 523-527, 2001
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11586360
 A:Accession: AF0498
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-206 <KUR>
 A:Cross-references: GB:AL590842; PIDN:CAC93555.1; PID:q15981995; GSPDB:GN00175
 C:Genetics:
 A:Gene: YPO4106

Query Match 33.3%; Score 5; DB 2; Length 206;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPES 12
 DB 76 STPES 80

RESULT 91

T17623
 hypothetical protein A133R - Chlorella virus PBCV-1
 C:Species: Chlorella virus PBCV-1
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T17623
 R:Graves, M.V.; Van Elten, J.L.
 submitted to the EMBL Data Library, May 1999
 A:Reference number: Z1806
 A:Accession: T17623
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA

A:Residues: 1-207 <GRA>
 A:Cross-references: EMBL:U42580; NID:g4028896; PIDN:AAC96501.1
 A:Experimental source: specific host Chlorella strain NC64A
 C:Genetics:
 A:Note: A133R

Query Match 33.3%; Score 5; DB 2; Length 207;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGP 5
 DB 108 SHLGP 112

RESULT 92

A48567
 calmodulin-ubiquitin associated gene CUB2.65 - Trypanosoma cruzi
 C:Species: Trypanosoma cruzi
 C>Date: 01-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 24-Sep-1999
 C:Accession: A48567
 R:Ajiloka, J.; Swindle, J.
 Mol. Biochem. Parasitol. 57, 127-136, 1993
 A:Title: The calmodulin-ubiquitin associated genes of Trypanosoma cruzi: their identifi
 A:Reference number: A48567; MUID:93149197; PMID:8381204
 A:Accession: A48567
 A:Status: preliminary
 A:Molecule type: nucleic acid
 A:Residues: 1-208 <AUJ>
 A:Cross-references: GB:L01583; NID:g162020; PIDN:AAA30171.1; PID:g162021
 A:Note: sequence extracted from NCBI backbone (NCBIN:123773; NCBI:123775)
 C:Superfamily: EF-hand protein EFH5

Query Match 33.3%; Score 5; DB 2; Length 208;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
 DB 170 RSTPE 174

RESULT 93

AB2157
 hypothetical protein alr2809 [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp.
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
 C:Accession: AB2157
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irligu
 Nakazaki, N.; Shimp, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata
 DNA Res. 8, 205-213, 2001
 A:Title: Complete genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AB2157
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-214 <KUR>
 A:Cross-references: GB:BA000019; PIDN:BA074508.1; PID:q17131902; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: alr2809
 C:Superfamily: conserved hypothetical protein s111186

Query Match 33.3%; Score 5; DB 2; Length 214;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRRA 15
 DB 35 ESRRA 39

A:Cross-references: DDBJ:AP000064; NID:95105945; PIDN:BAA81426.1; PID:d1045212; PID:9510
 A:Experimental source: strain K1
 C:Genetics:
 A:Gene: APE2411
 C:Superfamily: Aeropyrum pernix hypothetical protein APE2411

Query Match 33.3%; Score 5; DB 2; Length 197;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SHLGP 5
 Db 17 SHLGP 21

RESULT 85
 A45067
 laminin B1 chain variant - chicken (fragment)
 C:Species: Gallus gallus (chicken)
 C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 10-Dec-1999
 C:Accession: A45067
 R:O'Rear, J.J.
 J. Biol. Chem. 267, 20555-20557, 1992
 A:Title: A novel laminin B1 chain variant in avian eye.
 A:Reference number: A45067; MUID:93015947; PMID:1400373
 A:Accession: A45067
 A:Status: preliminary
 A:Molecule type: nucleic acid
 A:Residues: 1-198 <OLR>
 A:Cross-references: GB:L00963; NID:g212883; PIDN:AAA9140.1; PID:g212884
 A:Experimental source: eye
 A:Note: sequence extracted from NCBI backbone (NCBIP:115996)
 C:Superfamily: laminin beta-1 chain; laminin-type EGF-like homology
 C:Keywords: basement membrane; calcium binding; cell binding; coiled coil; disulfide bond
 F:82-87/Disulfide bonds: #status predicted

Query Match 33.3%; Score 5; DB 2; Length 198;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 10 PESRA 14
 Db 34 PESRA 38

RESULT 86
 A12631
 hypothetical protein Atu0452 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
 C:Species: Agrobacterium tumefaciens
 C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
 C:Accession: A12631
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
 erge, G.; Gallet, W.; Grant, C.; Guenther, D.; Kutyavlin, T.; Levy, R.; Li, M.; McClellan
 ; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A:Authors: 100, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
 ster, E.W.
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A:Reference number: AB2577; PMID:11743193
 A:Accession: A12631
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-201 <KUR>
 A:Cross-references: GB:AE008688; PIDN:AAL1471.1; PID:q17738796; GSPDB:GN00186
 C:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: Atu0452
 A:Map position: circular chromosome

Query Match 33.3%; Score 5; DB 2; Length 201;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRAA 15
 Db 163 ESRAA 167

RESULT 87
 B86488
 hypothetical protein F103.5 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
 C:Accession: B86488
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maitl, R.; Marzla
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Talli
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: B86488
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-201 <STO>
 A:Cross-references: GB:AE005172; NID:g11094738; PIDN:AA629672.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1

Query Match 33.3%; Score 5; DB 2; Length 201;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 LGPFR 7

Db 100 LGPFR 104

RESULT 88
 T07771
 probable heme transport protein - tomato mitochondrion
 C:Species: mitochondrion Lycopersicon esculentum (tomato)
 C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 20-Jun-2000
 C:Accession: T07771
 R:Shikanai, T.; Nakata, S.; Harada, K.; Watanabe, K.
 Plant Cell Physiol. 37, 692-696, 1996
 A:Title: Analysis of the heterologous transcripts of the highly edited orf206 in toma
 A:Reference number: Z16123; MUID:96416432; PMID:8819315
 C:Accession: T07771
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-206 <SHI>
 A:Cross-references: EMBL:D84426; PIDN:BA12352.1
 A:Experimental source: cultivar Sekai-Ichi; Leaves
 C:Genetics:
 A:Genome: mitochondrion
 C:Function:
 A:Description: involved in the biogenesis of cytochrome c
 C:Superfamily: cytochrome c biogenesis protein Cyw
 C:Keywords: mitochondrion; RNA editing

Query Match 33.3%; Score 5; DB 2; Length 206;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SHLGP 5

Db 46 SHLGP 50

RESULT 89
 T09503

RESULT 80
G90855
hypothetical protein ECs1815 [imported] - Escherichia coli (strain O157:H7, substrain R1
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: G90855
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: G90855
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-189 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA935238.1; PID:G13361280; GSPDB:GN00154
C:Genetics:
A:Gene: ECs1815

Query Match 33.3%; Score 5; DB 2; Length 189;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPES 12
|||||
Db 24 STPES 28

RESULT 81
A86369
hypothetical protein F508.9 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
C:Accession: A86369
R:Phellogis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizart, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzilli,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: A86369
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-189 <STO>
A:Cross-references: GB:AE005172; NID:G4056436; PIDN:ACG98009.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 33.3%; Score 5; DB 2; Length 189;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRAA 15
|||||
Db 68 ESRAA 72

RESULT 82
A82613
conserved hypothetical protein XF2007 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 03-Jun-2002
C:Accession: A82613
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347

A:Note: for a complete list of authors see reference number A59328 below
A:Accession: A82613
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-191 <SI>
A:Cross-references: GB:AB004019; GB:AE003849; NID:G9107105; PIDN:AA984809.1; GSPDB:GN
A:Experimental source: strain 9a5c
R:Rimpson, A.J.G.; Reinach, P.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.
Briões, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carre
as-Neto, E.; Docena, C.; El-Doriry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to Genbank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Frega, J.S.; Franco, S.C.; Franco, M.C.; Fr
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmitieri,
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Sanfeli, R.V.; Sava
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tsunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF2007
C:Superfamily: Methanobacterium thermoautotrophicum NADPH-oxidoreductase

Query Match 33.3%; Score 5; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HIGPH 6
|||||
Db 154 HIGPH 158

RESULT 83
T28682
hypothetical protein - Streptomyces coelicolor (fragment)
C:Species: Streptomyces coelicolor
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T28682
R:Parkhill, J.; Bentley, S.D.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, January 1999
A:Reference number: Z20512
A:Accession: T28682
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-191 <PAR>
A:Cross-references: EMBL:AL023496; NID:e1292348; PID:e1370577; PIDN:CAA18999.1

Query Match 33.3%; Score 5; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHR 7
|||||
Db 9 LGPHR 13

RESULT 84
B72471
hypothetical protein APE2411 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jun-2000
C:Accession: B72471
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Ta
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: B72471
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-197 <KAW>

RESULT 75

F75274

Conserved hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000

C:Accession: F75274

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

M.; Shen, M.; Yamathayan, P.J.; Lam, P.; McDonald, L.; Utterback, T.; Zaleski, C.; Ma

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

S:Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; M0ID:20036896; PMID:10567266

A:Accession: F75274

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-173 <WHI>

A:Cross-references: GB:AE002073; GB:AE000513; NID:g6460244; PIDN:AAF11973.1; PID:g646024

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR2432

A:Map position: 1

Query Match 33.3%; Score 5; DB 2; Length 173;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRRA 15
|||||
DB 80 ESRRA 84

RESULT 76

S14747

sphingomyelin phosphodiesterase (EC 3.1.4.12) - human (fragments)

C:Species: Homo sapiens (man)

C:Date: 21-Nov-1993 #sequence_revision 13-Mar-1997 #text_change 13-Mar-1997

C:Accession: S14747

R:Kurtz, J.; Stoffel, W.

BIOL. Chem. Hoppe-Seyler 372, 215-223, 1991

A:Title: Human placental sphingomyelinase. Purification to homogeneity, antigenic proper

A:Reference number: S14747; M0ID:91273814; PMID:2054099

A:Accession: S14747

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-16;17-32;33-43;44-57;58-74;75-84;85-99;100-107;108-127;128-138;139-153;154

C:Keywords: phosphoric diester hydrolase

Query Match 33.3%; Score 5; DB 2; Length 173;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPFR 7
|||||
DB 110 LGPFR 114

RESULT 77

T36394

probable pantoate-amino acid ligase - Streptomyces coelicolor (fragment)

C:Species: Streptomyces coelicolor

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Nov-2000

C:Accession: T36394

R:Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, April 1999

A:Reference number: Z21573

A:Accession: T36394

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-176 <OLI>

A:Cross-references: EMBL:AL049628; PIDN:CAB40883.1; GSPDB:GN00070; SCOEDB:SCE94.34c

A:Experimental source: strain A3(2)

C:Genetics:
A:Gene: SCOEDB:SCE94.34c
C:Superfamily: pantoate-beta-alanine ligase

Query Match 33.3%; Score 5; DB 2; Length 176;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRRA 15
|||||
DB 85 ESRRA 89

RESULT 78

A11820

hypothetical protein alr0113 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp.

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002

C:Accession: A11820

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kurlitz, T.; Sasamoto, S.; Watanabe, A.; Irigu

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium

A:Reference number: AB1807; M0ID:21595285; PMID:11759840

A:Accession: A11820

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-183 <KUR>

A:Cross-references: GB:BA000019; PIDN:BA077637.1; PID:g17135091; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

Query Match 33.3%; Score 5; DB 2; Length 183;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STRES 12
|||||
DB 9 STRES 13

RESULT 79

T01887

hypothetical protein F8M12.19 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 24-Mar-1999

C:Accession: T01887

R:Madsen, C.; Graves, T.; Cotton, M.; Modde, T.

submitted to the EMBL Data Library, July 1998

A:Description: The sequence of A. thaliana F8M12.

A:Reference number: Z14450

A:Accession: T01887

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-185 <MAD>

A:Cross-references: EMBL:AF080118; NID:g3513725; PID:g3513743

A:Experimental source: cultivar Columbia

C:Genetics:

A:Map position: 4

A:Introns: 64/3; 76/3

A:Note: F8M12.19

Query Match 33.3%; Score 5; DB 2; Length 185;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPFR 7
|||||
DB 126 LGPFR 130

Db 29 STPEs 33

RESULT 70

DB7241

VPS29-like phosphoesterase-related protein ML2654 [similarity] - Mycobacterium leprae
C:Species: Mycobacterium leprae
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 24-Aug-2001

C:Accession: D87241

R:Coile, S.T.; Eigmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001

A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sc
A:Title: Massive gene decay in the leprosy bacillus
A:Reference number: A86909; MUID:21128732; PMID:11234002

A:Accession: D87241

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-165 <STO>

A:Cross-references: GB:AL450380; NID:g13093861; PIDN:GAC32186.1; GSPDB:GN00147

C:Genetics:

A:Gene: ML2654

C:Superfamily: human vacuolar protein-sorting protein VPS29 homology; phosphoesterase co
F:2-62/Domain: phosphoesterase core homology <PEC>

Query Match

Best Local Similarity 33.3%; Score 5; DB 2; Length 165;
Pred. No. 1.6e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRAA 15

Db 48 ESRAA 52

RESULT 71

T16984

transcription factor homolog BRF3 - curled-leaved tobacco
C:Species: Nicotiana glauca (curled-leaved tobacco)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Jun-2000

C:Accession: T16984

R:Bot15Juk, N.V.

submitted to the EMBL Data Library, October 1996

A:Reference number: Z18621

A:Accession: T16984

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-165 <BOR>

A:Cross-references: EMBL:Y09106

A:Experimental source: somatic embryo

C:Superfamily: transcription factor BRF3

C:Keywords: transcription factor

Query Match

Best Local Similarity 33.3%; Score 5; DB 2; Length 165;
Pred. No. 1.6e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRA 14

Db 120 PESRA 124

RESULT 72

149694

glucokinase - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Jun-1999

C:Accession: 149694

R:Hughes, S.D.; Quade, C.; Milburn, J.L.; Cassidy, L.; Newgard, C.B.
J. Biol. Chem. 266, 4521-4530, 1991

A:Title: Expression of normal and novel glucokinase mRNAs in anterior pituitary and isle
A:Reference number: 149694; MUID:91154262; PMID:1999433

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-166 <RES>

A:Cross-references: GB:M5875; NID:g193538; PIDN:AAA37703.1; PID:g553920

C:Superfamily: hexokinase; hexokinase homology

F:25-166/Domain: hexokinase homology (fragment) <HXK>

Query Match

Best Local Similarity 33.3%; Score 5; DB 2; Length 166;
Pred. No. 1.6e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11

Db 63 RSTPE 67

RESULT 73

184740

glucokinase - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Jun-1999

C:Accession: I84740

R:Hughes, S.D.; Quade, C.; Milburn, J.L.; Cassidy, L.; Newgard, C.B.

J. Biol. Chem. 266, 4521-4530, 1991

A:Title: Expression of normal and novel glucokinase mRNAs in anterior pituitary and 1

A:Reference number: 149694; MUID:91154262; PMID:1999433

A:Accession: I84740

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-166 <RES>

A:Cross-references: GB:M5875; NID:g204371; PIDN:AAA41236.1; PID:g554439

C:Superfamily: hexokinase; hexokinase homology

F:25-166/Domain: hexokinase homology (fragment) <HXK>

Query Match

Best Local Similarity 33.3%; Score 5; DB 2; Length 166;
Pred. No. 1.6e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11

Db 63 RSTPE 67

RESULT 74

S68480

succinate dehydrogenase homolog YLR164w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein L9632.1

C:Species: Saccharomyces cerevisiae
C:Date: 20-Jul-1996 #sequence_revision 23-Aug-1996 #text_change 19-Apr-2002

C:Accession: S68480

R:Vaundin, M.

submitted to the EMBL Data Library, July 1996

A:Description: The sequence of S. cerevisiae cosmid 9362.

A:Reference number: S68471

A:Accession: S68480

A:Molecule type: DNA

A:Residues: 1-168 <VAU>

A:Cross-references: EMBL:U51921; NID:g1234842; PID:g1234843; GSPDB:GN00012; MIPS:YLR1

A:Gene: MIPS:YLR164w

A:Cross-references: SGD:S0004154

A:Map position: 12R

C:Keywords: transmembrane protein

F:66-82/Domain: transmembrane #status predicted <TM1>

F:124-140/Domain: transmembrane #status predicted <TM2>

Query Match

Best Local Similarity 33.3%; Score 5; DB 2; Length 168;
Pred. No. 1.7e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRA 14

Db 51 PESRA 55

```
RESULT 65
AD2345
hypothetical protein all4315 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp.
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AD2345
R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kunitz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MIMD:21595285; PMID:11759840
A:Accession: AD2345
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-145 <RUR>
A:Cross-references: GB:BA000019; PIDN:BAW6014.1; PID:g17133451; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all4315

Query Match
Best Local Similarity 100.0%; Score 5; DB 2; Length 145;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPES 12
Db 71 STPES 75

RESULT 66
T08734
hypothetical protein DKFZp566F0546.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 13-Aug-1999
C:Accession: T08734
R:Oltjenwalder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wilmann, S.
submitted to the Protein Sequence Database, May 1999
A:Reference number: Z16474
A:Accession: T08734
A:Molecule type: mRNA
A:Residues: 1-150 <OTT>
A:Cross-references: EMBL:AL050075
A:Experimental source: fetal kidney; clone DKFZp566F0546
C:Genetics:
A:Note: DKFZp566F0546.1

Query Match
Best Local Similarity 100.0%; Score 5; DB 2; Length 150;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHR 7
Db 36 LGPHR 40

RESULT 67
A97573
nitrogen regulatory protein P-II [imported] - Agrobacterium tumefaciens (strain C58, Cer
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 17-May-2002
C:Accession: A97573
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorllo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; PMID:11743194
A:Accession: A97573
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-157 <KUR>
```

```
A:Cross-references: GB:AE007869; PIDN:AAK87538.1; PID:g15156872; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C_3252
A:Map position: circular chromosome
C:Superfamily: regulatory protein P-II

Query Match
Best Local Similarity 100.0%; Score 5; DB 2; Length 157;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRA 14
Db 5 PESRA 9

RESULT 68
A71062
hypothetical protein PH191 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 20-Jun-2000
C:Accession: A71062
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Se
M.; Oniku, Y.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kushida, N.; Ogu
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
A:Reference number: A71000; MIMD:98344137; PMID:9679194
A:Accession: A71062
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-160 <KAW>
A:Cross-references: GB:AP000005; NID:g3236132; PIDN:BA30291.1; PID:g3257608
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by GenBa
C:Genetics:
A:Gene: PH191
C:Superfamily: Pyrococcus horikoshii hypothetical protein PH191

Query Match
Best Local Similarity 100.0%; Score 5; DB 2; Length 160;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 HRSTP 10
Db 4 HRSTP 8

RESULT 69
E84172
hypothetical protein Vng0121h [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: E84172
R:Ng, W.V.; Kennedy, S.P.; Mahaitas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Laaky
; Leithauser, B.; Keller, K.; Cruz, R.; Dawson, M.J.; Hough, D.W.; Maddocks, D.G.; Jé
Jung, K.H.; Alam, M.; Freilias, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.;
A:Title: Genome Sequence of Halobacterium species NRC-1.
A:Reference number: AB4160; MIMD:20504483; PMID:11016950
A:Accession: E84172
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-163 <STO>
A:Cross-references: GB:AE004437; NID:g10579769; PIDN:AAI8745.1; GSPDB:GN00138
C:Genetics:
A:Gene: VNG0121H

Query Match
Best Local Similarity 100.0%; Score 5; DB 2; Length 163;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPES 12
Db 11 STPES 12
```

A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-124 <A2E>
 A:Cross-references: EMBL:X76344; NID:9475914; PIDN:CA533963.1; PID:9475915
 C:Superfamily: human NADH dehydrogenase (ubiquinone) CI-B14 chain
 C:Keywords: NAD; oxidoreductase

Query Match 33.3%; Score 5; DB 2; Length 124;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
 |||||
 Db 34 RSTPE 38

RESULT 61

AH3491
 hypothetical cytosolic protein BME11918 [Imported] - Brucella melitensis (strain 16M)
 C:Species: Brucella melitensis
 C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002

A:Accession: AH3491
 R:DeVecchio, V.G.; Kaputral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, .; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letessier, Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
 A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
 A:Reference number: AD3252; PMID:11756688
 A:Accession: AH3491
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-127 <KUR>
 A:Cross-references: GB:AE008917; PIDN:AL53099.1; PID:g17983964; GSPDB:GN00190
 A:Experimental source: strain 16M
 C:Genetics:

A:Gene: BME11918
 A:Map position: 1

Query Match 33.3%; Score 5; DB 2; Length 127;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRAA 15
 |||||
 Db 46 ESRAA 50

RESULT 62

S23003
 trkA protein - Escherichia coli plasmid RP4
 C:Species: Escherichia coli
 C:Date: 31-Dec-1993 #sequence_revision 02-Aug-1994 #text_change 08-Oct-1999
 A:Accession: S23003

R:Ziegelin, G.; Ponsgrau, W.; Strack, B.; Balzer, D.; Kroege, M.; Kruff, V.; Lanka, E.
 DNA Seq. 1, 303-327, 1991
 A:Title: Nucleotide sequence and organization of genes flanking the transfer origin of R
 A:Reference number: S23292; MUID:92190548; PMID:1655997
 A:Accession: S23003
 A:Molecule type: DNA
 A:Residues: 1-134 <ZIE>
 A:Cross-references: EMBL:X54459; NID:942780; PIDN:CA53393.1; PID:942786
 A:Note: the authors did not translate the codon for residue 1
 C:Genetics:

A:Gene: trkA
 A:Superfamily: plasmid

Query Match 33.3%; Score 5; DB 2; Length 134;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRAA 15
 |||||
 Db 14 ESRAA 18

RESULT 63

C82821
 NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 3 [similarity] - Xylella fastidiosa
 C:Species: Xylella fastidiosa
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 03-Jun-2002
 A:Accession: C82821

R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
 Nature 406, 151-157, 2000
 A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A:Reference number: A82515; MUID:20365717; PMID:10910347
 A:Note: for a complete list of authors see reference number A59328 below
 A:Accession: C82821

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-135 <STM>
 A:Cross-references: GB:AE003884; GB:AE003849; NID:9105127; PIDN:AAE83116.1; GSPDB:GN
 A:Experimental source: strain 9a5c
 R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer
 as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
 submitted to GenBank, June 2000
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La
 chado, M.A.; Madeira, A.M.B.N.; Madeira, J.M.F.; Marino, C.L.; Marques, M.V.; Martins
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Palmieri,
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sava
 A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
 M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
 A:Reference number: A59328
 A:Contents: annotation
 C:Genetics:

A:Gene: XF0305
 C:Superfamily: NADH dehydrogenase (ubiquinone) chain 3
 C:Keywords: membrane-associated complex; NAD; oxidoreductase

Query Match 33.3%; Score 5; DB 2; Length 135;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHR 7
 |||||
 Db 46 LGPHR 50

RESULT 64

B84990
 50S ribosomal protein L16 [Imported] - Buchnera sp. (strain APS)
 C:Species: Buchnera sp.
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
 A:Accession: B84990

R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
 Nature 407, 81-86, 2000
 A:Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp
 A:Reference number: A84930; MUID:20445173; PMID:10993077
 A:Accession: B84990
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-136 <STO>
 A:Cross-references: GB:AP000398; GSPDB:GN00144
 A:Experimental source: strain APS
 C:Genetics:

A:Gene: rplP; BU517
 C:Superfamily: Escherichia coli ribosomal protein L16

Query Match 33.3%; Score 5; DB 2; Length 136;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRAA 15
 |||||
 Db 112 ESRAA 116

A:Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*
A:Reference number: AD3252; PMID:11756688
A:Accession: AD3436
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-87 <KUR>
A:Cross-references: GB:AE008917; PIDN:AL52655.1; PID:g17983479; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME11474
A:Map position: 1

Query Match 33.3%; Score 5; DB 2; Length 87;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRA 14
|||||
DB 77 PESRA 81

RESULT 56

C:ECG
C:cytochrome c [validated] - *Englena gracilis*
C:Species: *Englena gracilis*
C:Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 28-Jul-2000
C:Accession: A00068
R:Petigrew, G.W.; leaver, J.L.; Meyer, T.E.; Ryle, A.P.
Biochem. J. 147, 291-302, 1975
A:Title: Purification, properties and amino acid sequence of atypical cytochrome c from
A:Reference number: A00068; MUID:76039443; PMID:170910
A:Accession: A00068
A:Molecule type: protein
A:Residues: 1-102 <PEP>
C:Superfamily: cytochrome c; cytochrome c homology
C:Keywords: acetylated amino end; chromoprotein; electron transfer; heme; iron; metallo
F:4-97/Domain: cytochrome c homology <CYC>
F:1/Modified site: acetylated amino end (Gly) #status experimental
F:17/Inding site: heme (Cys) (covalent) #status experimental
F:18,79/Binding site: heme iron (His, Met) (axial ligands) #status predicted
F:85/Modified site: N6,N6,N6-trimethyllysine (Lys) #status experimental

Query Match 33.3%; Score 5; DB 1; Length 102;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRAA 15
|||||
DB 11 ESRAA 15

RESULT 57

AB0318
Conserved hypothetical protein YP02606 [Imported] - *Yersinia pestis* (strain CO92)
C:Species: *Yersinia pestis*
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
C:Accession: AB0318
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarrage, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,
Nature 413, 523-527, 2001
A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AB0318
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-105 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC92849.1; PID:g15980593; GSPDB:GN00175
C:Genetics:
A:Gene: YP02606
C:Superfamily: *Escherichia coli* ybeB protein

Query Match 33.3%; Score 5; DB 2; Length 105;

Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRAA 15
|||||
DB 58 ESRAA 62

RESULT 58

S70089
KORa protein - *Amycolatopsis methanolica*
C:Species: *Amycolatopsis methanolica*
C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 07-May-1999
C:Accession: S70089
R:Yrjöboed, J.W.; Jellinkova, M.; Hesses, G.I.; Dijkhuizen, L.
Mol. Microbiol. 18, 21-31, 1995
A:Title: Identification of the minimal replicon of plasmid pMEA300 of the methylotrophic
A:Reference number: S70087; MUID:96154938; PMID:8596458
A:Accession: S70089
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-118 <VR1>
A:Cross-references: EMBL:L36679
C:Genetics:
A:Gene: korA

Query Match 33.3%; Score 5; DB 2; Length 118;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESR 13
|||||
DB 76 TPESR 80

RESULT 59

A05114
Hypothetical protein E-122 - *Halobacterium salinarum* insertion sequence ISH50
C:Species: *Halobacterium salinarum*
C:Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 21-Jul-2000
C:Accession: A05114
R:Xu, W.L.; Doolittle, W.F.
Nucleic Acids Res. 11, 4195-4199, 1983
A:Reference number: A93475; MUID:83246542; PMID:6306577
A:Accession: A05114
A:Molecule type: DNA
A:Residues: 1-122 <XUM>
A:Cross-references: GB:X01584; NID:943515; PIDN:CAB37935.1; PID:el390961; PID:g446743
A:Note: the source is designated as *Halobacterium halobium*
C:Genetics:
A:Mobile element: insertion sequence ISH50
A:Start codon: GTG

Query Match 33.3%; Score 5; DB 2; Length 122;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESR 13
|||||
DB 4 TPESR 8

RESULT 60

S43840
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) - *Neurospora crassa*
C:Species: *Neurospora crassa*
C:Date: 13-Jan-1995 #sequence_revision 13-Mar-1997 #text_change 03-Jun-2002
C:Accession: S43840
R:Azevedo, J.E.; Eckerskorn, C.; Werner, S.
Biochem. J. 299, 297-302, 1994
A:Title: In organello assembly of respiratory-chain complex I: primary structure of t
A:Reference number: S43840; MUID:94220045; PMID:8166654
A:Accession: S43840

RC STRAIN-CV, SEKAI-ICHI, TISSUE-LEAF;
 RA Shikanei T., Nakata S., Harada K., Watanabe K.;
 RT "Analysis of the heterologous transcripts of the highly edited orf206
 in tomato mitochondria."
 RL Plant Cell Physiol. 0:0-0(1996).
 DR EMBL: D84426; BAA12352.1; -;
 DR InterPro: IPR003544; CytC_blog_CcMB.
 DR Pfam: PF03379; CcMB; 1.
 DR PRINTS: PR01414; CCMBBIOGNISIS.
 KW Mitochondrion.
 SQ SEQUENCE 206 AA; 23069 MM; 381F42431443295D CRC64;

Query Match 33.3%; Score 5; DB 8; Length 206;
 Best Local Similarity 100.0%; Pred. No. 5.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGP 5
 |||||
 DB 46 SHLGP 50

RESULT 200
 079373
 ID 079373 PRELIMINARY; PRT; 206 AA.
 AC 079373;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE He1B-homologue protein.
 GN CCB206.
 OS Brassica napus (Rape).
 OG Mitochondrion.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Brassica.
 OX NCBI_TaxID=3708;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, ISUZU-NATANE; TISSUE-LEAF;
 RA Itani K., Handa H.;
 RT "Repesed mitochondrial ccb206, a gene involved in cytochrome c
 biogenesis is cotranscribed with the nad3 and rps12 genes;
 RT organization, transcription, and RNA editing of the nad3/rps12/ccb206
 RT locus.";
 RL Curr. Genet. 0:0-0(1998).
 DR EMBL: D13697; BAA32558.1; -;
 DR InterPro: IPR003544; CytC_blog_CcMB.
 DR Pfam: PF03379; CcMB; 1.
 DR PRINTS: PR01414; CCMBBIOGNISIS.
 KW Mitochondrion.
 SQ SEQUENCE 206 AA; 23099 MM; 89A9C8F6F41D5134 CRC64;

Query Match 33.3%; Score 5; DB 8; Length 206;
 Best Local Similarity 100.0%; Pred. No. 5.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGP 5
 |||||
 DB 46 SHLGP 50

Search completed: March 10, 2003, 14:29:57
 Job time : 94 secs

THIS PAGE BLANK (USPTO)

CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, hematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activity/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX

Sequence 134 AA:

Query Match 40.0%; Score 6; DB 22; Length 134;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 PRRSTP 10
 |||||
 Db 51 PRRSTP 56

RESULT 108

ABG06201
 ID ABG06201 standard; Protein; 135 AA.

XX
 AC ABG06201;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #6192.

XX Human: chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PA 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

DR N-PSDB; AAS70388.

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

XX Claim 20: SEQ ID NO 36560; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.

CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX

Sequence 135 AA:

Query Match 40.0%; Score 6; DB 22; Length 135;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHRS 8
 |||||
 Db 55 LGPHRS 60

RESULT 109

ABG14117
 ID ABG14117 standard; Protein; 135 AA.

XX
 AC ABG14117;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #14108.

XX Human: chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PA 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

DR N-PSDB; AAS78304.

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

XX Claim 20: SEQ ID NO 44476; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

SQL Sequence 135 AA;
Query Match 40.0%; Score 6; DB 22; Length 135;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHRS 8
|||||
Db 55 LGPHRS 60

RESULT 110
ABG14498
ID ABG14498 standard; Protein: 137 AA.
XX
AC ABG14498;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #14489.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR MPI: 2001-639362/73.
DR N-PSDB: AAS78685.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 44857; 103bp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

SQL Sequence 137 AA;
Query Match 40.0%; Score 6; DB 22; Length 137;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHRS 8
|||||
Db 54 LGPHRS 59

RESULT 111
ABG11461
ID ABG11461 standard; Protein: 146 AA.
XX
AC ABG11461;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #11452.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR MPI: 2001-639362/73.
DR N-PSDB: AAS75648.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 41820; 103bp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 146 AA;

Query Match 40.0%; Score 6; DB 22; Length 146;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHRS 8
| | | | |
Db 85 LGPHRS 90

RESULT 112

ABG28247
ID ABG28247 standard; Protein: 158 AA.

XX AC ABG28247;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #28238.

XX Human: chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

DR N-PSDB; AAS92434.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

PS Claim 20: SEQ ID No 58606; 103pp: English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 158 AA;

Query Match 40.0%; Score 6; DB 22; Length 158;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHRS 8
| | | | |
Db 63 LGPHRS 68

RESULT 113

ABG06433
ID ABG06433 standard; Protein: 161 AA.

XX AC ABG06433;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #6424.

XX Human: chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

DR N-PSDB; AAS70620.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

PS Claim 20: SEQ ID No 36792; 103pp: English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 161 AA;

Query Match 40.0%; Score 6; DB 22; Length 161;

Best Local Similarity 100.0%; Pred. No. 55;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPHRS 8

DB 103 LGPHRS 108

RESULT 114

ABG19252

ID ABG19252 standard; Protein; 167 AA.

AC ABG19252;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #19243.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Dmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR N-PSDB; AAS83439.

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity

PS Claim 20; SEQ ID No 49611; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 167 AA;

Query Match 40.0%; Score 6; DB 22; Length 167;

Best Local Similarity 100.0%; Pred. No. 57;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPHRS 8

DB 55 LGPHRS 60

RESULT 115

ABG14481

ID ABG14481 standard; Protein; 176 AA.

AC ABG14481;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #14472.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Dmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR N-PSDB; AAS78668.

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity

PS Claim 20; SEQ ID No 44840; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 176 AA;

Query Match

Best Local Similarity 40.0%; Score 6; DB 22; Length 176;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 LGPHRS 8
|||||
Db 118 LGPHRS 123

RESULT 116

ABG06238
ID ABG06238 standard; Protein: 177 AA.

AC ABG06238;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #6229.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PE 30-MAR-2001: 2001WO-US08631.

PR 31-MAR-2000: 2000US-0540217.

PR 23-AUG-2000: 2000US-0649167.

PA (HYSE-) HYSEQ INC

PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

DR N-PSDB: AAS70425.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

PS Claim 20: SEQ ID NO 36597; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 177 AA;

Query Match 40.0%; Score 6; DB 22; Length 177;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPHRS 8
|||||
Db 55 LGPHRS 60

RESULT 117

ABG28263
ID ABG28263 standard; Protein: 199 AA.

AC ABG28263;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #28254.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PE 30-MAR-2001: 2001WO-US08631.

PR 31-MAR-2000: 2000US-0540217.

PR 23-AUG-2000: 2000US-0649167.

PA (HYSE-) HYSEQ INC

PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

DR N-PSDB: AAS92450.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

PS Claim 20: SEQ ID NO 58622; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 199 AA;

Query Match 40.0%; Score 6; DB 22; Length 199;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 LGPHRS 8
|||||

Db 88 LGPHRS 93

RESULT 118

ID AAB65725 standard; Protein; 205 AA.

AC AAB65725;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #28248.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW Food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR N-PSDB; AAS92444.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

PS Claim 20; SEQ ID NO 58616; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAG00010-ABG30377 represent novel human CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed CC specification, but was obtained in electronic format directly from WIPO CC at ftp.wipo.int/pub/published_pcl_sequences.

SQ Sequence 205 AA;

Query Match

Best Local Similarity 40.0%; Score 6; DB 22; Length 205;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHRS 8

Db 181 LGPHRS 186.

RESULT 119

ID AAB65725 standard; Protein; 256 AA.

AC AAB65725;

DT 27-MAR-2001 (first entry)

DE Lethal leaf spot protein 11s1-related protein #2.

KW Cell death modulator; programmed cell death; PCD; apoptosis;
KW forestry plant.

OS Pinus radiata.

PN WO200075331-A1.

PD 14-DEC-2000.

PF 02-JUN-2000; 2000WO-NZ00086.

PR 04-JUN-1999; 99US-0325932.

PA (GENE-) GENESIS RES & DEV CORP LTD.
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.

PI Flinn B, Lasham A;

DR WPI; 2001-061724/07.

DR N-PSDB; AAF44751.

PT Novel defender against cell death polynucleotide useful for modulating
PT programmed cell death pathway and specific development pathways in
PT forestry plant.

PS Claim 22; Pages 67-68; 142pp; English.

The present invention relates to coding sequences (see AAF44740-F44840 CC and AAF44843-F44844) and proteins (see AAB65714-B65814) involved in programmed cell death (PCD; apoptosis). The coding sequences and proteins of the present invention are useful for modulating a PCD or cell death CC pathway and various developmental pathways in a forestry plant, by CC stably incorporating one of the present coding sequences into the genome CC of the forestry plant, where the coding sequence provides a PCD pathway CC that is not present in a native form of the forestry plant.

SQ Sequence 256 AA;

Query Match

Best Local Similarity 40.0%; Score 6; DB 22; Length 256;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRAA 15

Db 184 PESRAA 189

RESULT 120

ID AAY43977 standard; Protein; 270 AA.

AC AAY43977;

DT 21-DEC-1999 (first entry)

DE Mouse protein kinase #7.

KW Prediction; secondary structure; alignment; evolutionary conservation;
KW homology; periodicity; co-variation analysis; antigenic site;
KW site directed mutagenesis; interaction.

OS Mus sp.

XX US958784 A.
PN 28-SEP-1999
XX 25-MAR-1992; 92US-0857224.
XX 25-MAR-1992; 92US-0857224.
XX (BENNY) BENNER S A.
XX Benner SA;
XX MPI; 1999-570766/48.
XX Predicting the folded structure of proteins -
XX Disclosure; Column 305-308; 113pp; English.
XX Sequences AA43902-Y44015 represent proteins used in a novel method of
CC predicting the folded structure of proteins, by aligning sequences of
CC homologous proteins and using patterns of evolutionarily conserved and
CC varied sequences to assign positions. Positions in the alignment are
CC assigned to the surface or inside of the folded structure, active sites,
CC and pairing segments. Secondary structural units are assigned by
CC identifying periodicity in the assignments, and assembled into globular
CC form using distance constraints imposed by disulfide bridges, active
CC site assignments and co-variation analysis. The predicted secondary
CC structures are useful for identifying antigenic sites on a protein
CC molecule, as guides for site directed mutagenesis studies, and for
CC understanding the interaction of a protein with other molecules.
SQ Sequence 270 AA;

Query Match 40.0%; Score 6; DB 20; Length 270;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHLGP 6
| | | | |
DB 59 SHLGP 64

RESULT 121
ID ABO8486 standard; Protein; 303 AA.
AC ABO8486;
XX 13-FEB-2002 (first entry)
DE Novel human diagnostic protein #8477.
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX Homo sapiens.
OS WO200175067-A2.
PN 11-OCT-2001.
PD 30-MAR-2001; 2001MO-US08631.
PF 31-MAR-2001; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSE INC.
XX Drmanac RT, Liu C, Tang YF.
XX MPI; 2001-639362/73.
DR N-PSDB; AAS72673.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX Claim 20; SEQ ID No 38845; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABO00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at http://ipo.int/pub/published_pct_sequences.
SQ Sequence 303 AA;

Query Match 40.0%; Score 6; DB 22; Length 303;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5 PHRSTP 10
| | | | |
DB 296 PHRSTP 301

RESULT 122
ID AAW24231 standard; Protein; 329 AA.
AC AAW24231;
XX 17-MAR-1998 (first entry)
DE Human melanoma associated delayed early response splice variant.
KW Melanoma associated delayed early response gene; MADER gene;
KW MADER protein; growth alteration; malignant melanoma; breast carcinoma;
KW cancerous condition; MADER translocation event; MADER immunogen;
KW MADER antigen.
XX Homo sapiens.
OS WO9728193-A1.
PN 07-AUG-1997.
PD 30-JAN-1997; 97MO-US01586.
PF 30-JAN-1997; 97US-0593563.
PR 30-JAN-1997; 96US-0593563.
XX (MELC-) MELCOR DIAGNOSTICS INC.
XX Johnson JP;
XX MPI; 1997-402557/37.

PT Monoclonal antibody which binds to the MADER protein - used for
 XX detecting cancerous conditions, such as breast carcinoma
 PS Disclosure; Fig 3; 69pp; English.
 XX
 CC The present sequence represents a novel melanoma associated delayed
 CC early response (MADER) protein. This nuclear protein is associated
 CC with growth alterations in malignant melanomas and other cancerous
 CC conditions, and is over-expressed in human malignant melanomas. Several
 CC variants of the protein have been identified (AAW24228-31), the present
 CC sequence being short splice variant. The DNA encoding the present
 CC sequence has a deleted internal sequence, producing a frameshift
 CC that causes premature termination of translation. Termination results in
 CC the loss of approximately one-third of the full length 525 amino acid
 CC MADER molecule. Analysis of human placental DNA has revealed that this
 CC alternatively spliced form of MADER is found in a 1:1 ratio with the
 CC full length version. The shorter MADER has an altered function and fails
 CC to repress erg-1 activity. Chromosomal rearrangement of MADER can be
 CC detected by hybridising immobilised chromosomal target DNA, that has
 CC been rendered single stranded and is obtained from a cell suspected of
 CC having undergone a MADER translocation event with a single stranded
 CC oligonucleotide probe complementary to a MADER nucleotide sequence. The
 CC probe contains a moiety capable of direct or indirect visualisation.
 CC Antibodies raised against the MADER protein can be used for detecting a
 CC cancerous condition, particularly melanoma malignancies, and especially
 CC a breast carcinoma. A composition comprising a MADER immunogen and a
 CC pharmaceutically acceptable vehicle can be used to elicit an immune
 CC response against a cell which over-expresses a MADER antigen.
 CC
 SQ Sequence 329 AA:
 Query Match 40.0%; Score 6; DB 18; Length 329;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 7 RSTPES 12
 |||||
 Db 188 RSTPES 193
 RESULT 123
 AAW31903
 ID AAW31903 standard; Protein; 369 AA.
 AC AAW31903;
 XX
 DT 28-APR-1998 (first entry)
 XX
 DE Streptococcus pneumoniae histidyl tRNA synthetase.
 XX
 KM Histidyl tRNA synthetase; hss polypeptide; genetic immunisation;
 KM vaccine; antibacterial; antibiotic; otitis media; conjunctivitis;
 KM pneumonia; bacteraemia; meningitis; sinusitis; pleural empysema;
 KM endocarditis; gene therapy.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN WO9739017-A1.
 PD 23-OCT-1997.
 XX
 PF 18-APR-1997; 97WO-US06877.
 XX
 PR 18-APR-1996; 96GB-0007993.
 XX
 PA (SMK) SMITHKLINE BEECHAM CORP.
 PA (SMK) SMITHKLINE BEECHAM PLC.
 XX
 PI Lawlor EJ;
 DR WPI; 1997-526393/48.
 DR N-PSDB; AAT188770.
 XX

PT DNA encoding histidyl-tRNA synthetase from Streptococcus pneumoniae
 PT - useful for protection against bacterial infections.
 XX
 PS Claim 12; Page 33-34; 43pp; English.
 XX
 CC The present sequence represents the histidyl tRNA synthetase (hss) from
 CC Streptococcus pneumoniae 0100993 (NCIMB 40800). The hss polypeptides,
 CC antagonists, antibodies and related nucleic acids can be used for
 CC diagnosis and treatment of bacterial diseases. In particular, they are
 CC directed towards Streptococcus pneumoniae infections causing otitis
 CC media, conjunctivitis, pneumonia, bacteraemia, meningitis, sinusitis,
 CC pleural empysema and endocarditis. Hss polypeptides, or vectors for
 CC their expression, can be used prophylactically in vaccines to raise an
 CC antibody and/or a cell immune response against these same diseases.
 CC Additionally, the new polypeptides allow agonists and antagonists of
 CC hss to be identified using standard binding assays. The compounds which
 CC are identified may have useful bacteriostatic and/or bacteriocidal
 CC activity. Antibodies against hss can used to treat infection, isolate or
 CC identify hss expressing clones, purify hss and as an immunoassay
 CC reagent. More generally, the products can prevent adhesion of bacteria to
 CC wounds and in-dwelling devices, block hss protein mediated invasion of
 CC mammalian cells and block the normal progression of infection.
 CC
 SQ Sequence 369 AA:
 Query Match 40.0%; Score 6; DB 18; Length 369;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 10 PESRAA 15
 |||||
 Db 171 PESRAA 176
 RESULT 124
 AAW24229
 ID AAW24229 standard; Protein; 411 AA.
 AC AAW24229;
 XX
 DT 17-MAR-1998 (first entry)
 XX
 DE Human melanoma associated delayed early response Drop8 variant protein.
 XX
 KM Melanoma associated delayed early response gene; MADER gene;
 KM MADER protein; growth alteration; malignant melanoma; breast carcinoma;
 KM cancerous condition; MADER translocation event; MADER immunogen;
 KM MADER antigen.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FH Region 246..260
 FT /note="bipartite nuclear localisation signal"
 FT
 PN WO9728193-A1.
 PD 07-AUG-1997.
 XX
 PF 30-JAN-1997; 97WO-US01586.
 XX
 PR 30-JAN-1997; 97US-0593563.
 XX
 PR 30-JAN-1996; 96US-0593563.
 XX
 PA (MELC-) MELCORP DIAGNOSTICS INC.
 XX
 PI Johnson JP;
 DR WPI; 1997-402557/37.
 DR N-PSDB; AAT77839.
 XX
 PT Monoclonal antibody which binds to the MADER protein - used for
 PT detecting cancerous conditions, such as breast carcinoma

XX Claim 32: Page -: 69pp; English.
 PS
 CC The present sequence represents a novel melanoma associated delayed
 CC early response (MADR) protein. This 55 kDa nuclear protein is associated
 CC with growth alterations in malignant melanomas and other cancers
 CC conditions, and is over-expressed in human malignant melanomas. Several
 CC variants of the protein have been identified (AAW24228-31), the present
 CC protein being the Drop8 splice variant. The protein contains repeat
 CC motifs which are characteristic of gene regulatory DNA binding proteins.
 CC Chromosomal rearrangement of MADR can be detected by hybridizing
 CC immobilised chromosomal target DNA, that has been rendered single
 CC stranded and is obtained from a cell suspected of having undergone a
 CC MADR translocation event with a single stranded oligonucleotide probe
 CC complementary to a MADR nucleotide sequence. The probe contains a moiety
 CC capable of direct or indirect visualisation. Antibodies raised against
 CC the MADR protein can be used for detecting a cancerous condition,
 CC particularly melanoma malignancies, and especially a breast carcinoma. A
 CC composition comprising a MADR immunogen and a pharmaceutically
 CC acceptable vehicle can be used to elicit an immune response against a
 CC cell which over-expresses a MADR antigen.
 CC note: the present sequence is not given in the specification; it was
 CC created using information provided.
 XX
 SQ Sequence 411 AA;
 Query Match 40.0%; Score 6; DB 18; Length 411;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 7 RSTPES 12
 |||||
 DB 139 RSTPES 144
 RESULT 125
 AAU27706
 ID AAU27706 standard; Protein; 415 AA.
 AC AAU27706;
 XX
 DT 18-DEC-2001 (first entry)
 DE Human full-length polypeptide sequence #31.
 XX
 KW Mammal: human; rhesus monkey; baker's yeast; fission yeast; Norway rat;
 KW mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukemia;
 KW cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;
 KW nervous system disorder; inflammatory disorder; cell differentiation;
 KW angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;
 KW genetic disorder; bone regeneration; tendon; ligament; tissue repair;
 KW cytostatic; antineumatic; antiarthritic; vulnereary; antiinflammatory;
 KW antibacterial; immunosuppressive; vasotropic; antiParkinsonian;
 KW neuroprotective; osteopathic; antidiabetic; antiallergic;
 KW immunostimulant; analgesic; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO200164834-A2.
 XX
 PD 07-SEP-2001.
 PF 26-FEB-2001; 2001MO-US04926.
 XX
 PR 28-FEB-2000; 2000US-0515126.
 PR 18-MAY-2000; 2000US-0577409.
 PR 17-JUN-2000; 2000US-0597707.
 PR 14-JUL-2000; 2000US-0616807.
 PR 19-SEP-2000; 2000US-0664641.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;

PI Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;
 PI Drmanac R;
 XX
 DR WPI; 2001-589862/66.
 XX
 DR N-PSDB; AAS44606.
 XX
 PT Novel polypeptides and nucleic acids obtained from cDNA libraries
 PT prepared from various human tissues, for diagnosis, treatment of
 PT cancer, neurological, inflammatory disorders and for use in arrays for
 PT detection
 XX
 PS Claim 10: SEQ ID NO 203; 153pp; English.
 XX
 CC Sequences AAU27676-AAU28019 represent full-length polypeptides and
 CC contig polypeptides of the invention. The proteins and their associated
 CC DNA sequences are useful for the treatment, diagnosis and prevention of
 CC various types of disorder in a mammalian subject such as a human, dog,
 CC monkey, mouse, hamster or rat. The disorders include cancers such as
 CC leukemia, lymphoma and neuroblastoma, autoimmune disorders such as
 CC multiple sclerosis, connective tissue disease, rheumatoid arthritis,
 CC diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system
 CC disorders such as Parkinson's disease, Alzheimer's disease, Huntington's
 CC chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and
 CC Weirnicke disease, inflammatory disorders such as nephritis, Crohn's
 CC disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory
 CC bowel disease. The sequences exhibit activity relating to angiogenesis,
 CC cell proliferation, cell differentiation, stem cell growth factor,
 CC activin or inhibin. Therefore, they can be used to manipulate stem cells
 CC in culture to give rise to neuroepithelial cells that can be used to
 CC augment or replace cells damaged by illness, accidental damage or genetic
 CC disorders. The sequences may also be used for regeneration of bone,
 CC cartilage, tendons and ligaments and in tissue repair and burn healing.
 CC Note: Some sequences for this patent did not form part of the printed
 CC specification, but were obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 415 AA;
 Query Match 40.0%; Score 6; DB 22; Length 415;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 7 RSTPES 12
 |||||
 DB 31 RSTPES 36
 RESULT 126
 ABP27059
 ID ABP27059 standard; Protein; 426 AA.
 AC ABP27059;
 XX
 DT 02-JUL-2002 (first entry)
 DE Streptococcus polypeptide SEQ ID NO 3294.
 XX
 KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
 XX
 OS Streptococcus pyogenes.
 XX
 PN WO200234771-A2.
 XX
 PD 02-MAY-2002.
 PF 29-OCT-2004; 2001MO-GB04789.
 XX
 PR 27-OCT-2000; 2000GB-0026333.
 PR 24-NOV-2000; 2000GB-0028727.
 PR 07-MAR-2001; 2001GB-0005640.
 XX

PA (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Telford J, Masiann V, Margarit Ros VI, Grandi G, Fraser C;
 PI Tettein H;
 XX
 DR WPI: 2002-352536/38.
 DR N-PSDB: ABN67690.
 XX
 PT New Streptococcus protein for the treatment or prevention of infection
 PT or disease caused by Streptococcus bacteria, such as meningitis, and
 PT for detecting a compound that binds to the protein.
 XX
 PS Claim 1: Page 3483; 4525pp; English.
 XX
 CC The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GHS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins.
 XX
 SQ Sequence 426 AA:
 XX
 Query Match 40.0%; Score 6; DB 23; Length 426;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 10 PESRAA 15
 XX
 DB 171 PESRAA 176
 XX
 RESULT 127
 AAW70983
 ID AAW70983 standard; Protein; 429 AA.
 XX
 AC AAW70983;
 XX
 DT 16-OCT-1998 (first entry)
 XX
 DE Histidyl tRNA synthetase of Streptococcus pneumoniae.
 XX
 KW Histidyl tRNA synthetase; hists; S. equisimilis; diagnosis; disease;
 KW infection; Helicobacter pylori.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN US5795758-A.
 XX
 PD 18-AUG-1998.
 XX
 PF 06-AUG-1997; 97US-0906744.
 XX
 PR 06-AUG-1997; 97US-0906744.
 PR 18-APR-1997; 97US-0844055.
 XX
 PA (SMK) SMITHKLINE BEECHAM CORP.
 XX
 PI Gentry DR, Greenwood RC, Lawlor EJ;
 XX
 DR WPI: 1998-466670/40.

DR N-PSDB: AAW42887.
 XX
 PT Nucleic acid encoding Streptococcus pneumoniae histidyl tRNA
 PT synthetase - useful for recombinant production of the enzyme in
 PT diagnosis, treatment and prevention of Streptococcus infections, and
 PT for screening of inhibitors
 XX
 PS Claim 1: Columns 25-28; 16pp; English.
 XX
 CC The present sequence represents a Streptococcus pneumoniae histidyl tRNA
 CC synthetase (hists). The protein has homology to a S. equisimilis hists
 CC protein. Fragments of hists that retain binding and/or catalytic
 CC properties are used for research as novel targets against S. pneumoniae.
 CC The hists nucleic acid sequence (and primers and probes derived from it)
 CC are used to isolate related genes, to diagnose disease, specifically
 CC S. pneumoniae infection (e.g. otitis media, pneumonia, conjunctivitis,
 CC osteomyelitis or especially meningitis, possibly also infections by
 CC Helicobacter pylori and associated tumours), to identify mutations or
 CC polymorphisms for serotyping, and antisense sequences are potential
 CC hists antagonists.
 XX
 SQ Sequence 429 AA:
 XX
 Query Match 40.0%; Score 6; DB 19; Length 429;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 10 PESRAA 15
 XX
 DB 171 PESRAA 176
 XX
 RESULT 128
 AAW85136
 ID AAW85136 standard; Protein; 429 AA.
 XX
 AC AAW85136;
 XX
 DT 20-JUN-2000 (first entry)
 XX
 DE Histidyl tRNA synthetase (hists) amino acid sequence.
 XX
 KW Histidyl tRNA synthetase; protein synthesis; otitis media; pneumonia;
 KW conjunctivitis; bacteraemia; meningitis; sinusitis; pleural empyema;
 KW endocarditis; hists.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN US6040162-A.
 XX
 PD 21-MAR-2000.
 XX
 PF 08-JUN-1998; 98US-0093134.
 XX
 PR 06-AUG-1997; 97US-0906744.
 PR 18-APR-1997; 97US-0844055.
 XX
 PA (SMK) SMITHKLINE BEECHAM CORP.
 XX
 PI Lawlor EJ, Gentry DR, Greenwood RC;
 XX
 DR WPI: 2000-270137/23.
 DR N-PSDB: AA298858.
 XX
 PT Histidyl tRNA synthetase enzyme from streptococcus genus useful for
 PT treating diseases such as otitis media, conjunctivitis, pneumonia,
 PT bacteraemia, meningitis, sinusitis, pleural empyema and endocarditis
 XX
 PS Claim 1: Column 7-8; 16pp; English.
 XX
 CC This sequence represents a Streptococcus pneumoniae histidyl tRNA
 CC synthetase (hists) amino acid sequence. RNA synthetases have
 CC a primary role in protein synthesis, and inhibitors of bacterial RNA

CC synthetase have the potential to be antibacterial agents. The histidyl
 CC tRNA synthetase polypeptide sequence of the invention is useful for
 CC treating diseases such as otitis media, conjunctivitis, pneumonia,
 CC bacteremia, meningitis, sinusitis, pleural empyema and endocarditis. The
 CC protein can also be used to treat infections of the cerebrospinal fluid.
 CC The nucleotide sequence encoding the histidyl tRNA synthetase enzyme is
 CC useful for therapeutic and prophylactic treatments, particularly for
 CC genetic immunisation.

XX Sequence 429 AA;

Query Match 40.0%; Score 6; DB 21; Length 429;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRAA 15
 |||||
 DB 171 PESRAA 176

RESULT 129

AAU37913
 ID AAU37913 standard; Protein; 429 AA.

XX AC AAU37913;

DT 14-FEB-2002 (first entry)

DE Streptococcus pneumoniae cellular proliferation protein #342.

XX Antisense: prokaryotic cellular proliferation protein;

KM antibiotic; antibacterial; drug design.

XX Streptococcus pneumoniae.

XX WO200170955-A2.

PD 27-SEP-2001.

PF 21-MAR-2001; 2001WO-US09180.

XX 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207272P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;

DR N-PSDB; AAS55772.

XX WPI: 2001-611495/70.

XX N-PSDB; AAS55772.

XX Example 3; Seq ID No 13506; 511pp; English.

CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery

CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pcl_sequences.

XX Sequence 429 AA;

Query Match 40.0%; Score 6; DB 22; Length 429;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRAA 15
 |||||
 DB 171 PESRAA 176

RESULT 130

AAW24228
 ID AAW24228 standard; Protein; 475 AA.

XX AC AAW24228;

DT 17-MAR-1998 (first entry)

DE Human melanoma associated delayed early response Drop9 variant protein.

XX Melanoma associated delayed early response gene; MADER gene;

KM MADER protein; growth alteration; malignant melanoma; breast carcinoma;
 KW cancerous condition; MADER translocation event; MADER immunogen;

XX MADER antigen.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 237 /note= "putative N-linked glycosylation site"

FT Region 246..260 /note= "bipartite nuclear localisation signal"

FT Region 61..185 /note= "N-terminal region"

XX WO9728193-A1.

XX 07-APR-1997.

PF 30-JAN-1997; 97WO-US01586.

PR 30-JAN-1997; 97US-0593563.

PR 30-JAN-1996; 96US-0593563.

XX (MELC-) MELCORP DIAGNOSTICS INC.

XX Johnson JP;

DR N-PSDB; AAT77838.

XX WPI: 1997-402557/37.

XX Claim 31; Fig 1; 69pp; English.

CC Monoclonal antibody which binds to the MADER protein - used for
 CC detecting cancerous conditions, such as breast carcinoma
 CC The present sequence represents a novel melanoma associated delayed
 CC early response (MADER) protein. This 55 kDa nuclear protein is associated
 CC with growth alterations in malignant melanomas and other cancerous
 CC conditions, and is over-expressed in human malignant melanomas. Several
 CC variants of the protein have been identified (AAW24229-31), the present
 CC protein being the Drop9 variant. The protein contains repeat motifs which
 CC are characteristic of gene regulatory DNA binding proteins. Chromosomal

CC rearrangement of MADER can be detected by hybridising immobilised
CC chromosomal target DNA, that has been rendered single stranded and is
CC obtained from a cell suspected of having undergone a MADER translocation
CC event with a single stranded oligonucleotide probe complementary to a
CC MADER nucleotide sequence. The probe contains a moiety capable of direct
CC or indirect visualisation. Antibodies raised against the MADER protein
CC can be used for detecting a cancerous condition, particularly melanoma
CC malignancies, and especially a breast carcinoma. A composition
CC comprising a MADER immunogen and a pharmaceutically acceptable vehicle
CC can be used to elicit an immune response against a cell which
CC over-expresses a MADER antigen.

SO Sequence 475 AA;

Query Match 40.0%; Score 6; DB 18; Length 475;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPES 12
Db 139 RSTPES 144

RESULT 131
ABB97204
ID ABB97204 standard; Protein; 479 AA.
AC ABB97204;
XX
XX
DT 28-JUN-2002 (first entry)
XX
DE Novel human protein SEQ ID NO: 472.

XX Human: antihaemic; vulnary; antiinflammatory; immunomodulator;
XX antifertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
XX neuroprotective; antiparkinsonian; protein therapy; ESR;
XX expressed sequence tag.

OS Homo sapiens.
XX
XX WO20022660-A2.
PN
XX
XX 21-MAR-2002.
PD
XX
XX 10-SEP-2001; 9001WO-US26015.
PF
XX
XX 11-SEP-2000; 2000US-0659671.
PR
XX
XX (HYSE-) HYSEQ IND.

XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
XX Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
XX MPI: 2002-292408/73.
DR N-PSDB; ABN32390.
XX
XX An isolated polynucleotide for treating diseases associated with its
XX encoded polypeptide such as cancer and multiple sclerosis -
XX
XX Example 2; SEQ ID NO 472; 509pp; English.

XX The present invention provides the protein and coding sequences of 444
XX novel human proteins. These were isolated from expressed sequences tags
XX (ESTs). They can be used to stimulate cell growth, to regulate
XX haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
XX e.g. in burn treatment, to regulate the immune system e.g. to treat
XX multiple sclerosis, to regulate activin or inhibin e.g. to treat
XX infertility, to regulate haemostasis or thrombolysis e.g. to treat
XX stroke and cancer, to screen for drugs, to treat inflammatory conditions
XX e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
XX Parkinson's disease. The present sequence is a protein of the invention.

SO Sequence 479 AA;

Query Match 40.0%; Score 6; DB 23; Length 479;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPES 12
Db 95 RSTPES 100

RESULT 132
AAG34332
ID AAG34332 standard; Protein; 513 AA.
AC AAG34332;
XX
XX
DT 18-OCT-2000 (first entry)
XX
XX Zea mays protein fragment SEQ ID NO: 41755.
DE
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence; corn.

OS Zea mays subsp. mays.
XX
XX EP103405-A2.
PN
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX
XX 25-FEB-1999; 99US-0121825.
PR
XX
XX 05-MAR-1999; 99US-0123180.
PR
XX
XX 09-MAR-1999; 99US-0123548.
PR
XX
XX 23-MAR-1999; 99US-0125788.
PR
XX
XX 25-MAR-1999; 99US-0126264.
PR
XX
XX 29-MAR-1999; 99US-0126785.
PR
XX
XX 01-APR-1999; 99US-0127462.
PR
XX
XX 06-APR-1999; 99US-0128234.
PR
XX
XX 08-APR-1999; 99US-0128714.
PR
XX
XX 16-APR-1999; 99US-0129845.
PR
XX
XX 19-APR-1999; 99US-0130077.
PR
XX
XX 21-APR-1999; 99US-0130449.
PR
XX
XX 23-APR-1999; 99US-0130510.
PR
XX
XX 28-APR-1999; 99US-0130891.
PR
XX
XX 30-APR-1999; 99US-0131449.
PR
XX
XX 30-APR-1999; 99US-0132048.
PR
XX
XX 30-APR-1999; 99US-0132407.
PR
XX
XX 04-MAY-1999; 99US-0132484.
PR
XX
XX 05-MAY-1999; 99US-0132485.
PR
XX
XX 06-MAY-1999; 99US-0132486.
PR
XX
XX 06-MAY-1999; 99US-0132487.
PR
XX
XX 07-MAY-1999; 99US-0132863.
PR
XX
XX 11-MAY-1999; 99US-0134256.
PR
XX
XX 14-MAY-1999; 99US-0134218.
PR
XX
XX 14-MAY-1999; 99US-0134221.
PR
XX
XX 14-MAY-1999; 99US-0134370.
PR
XX
XX 18-MAY-1999; 99US-0134768.
PR
XX
XX 19-MAY-1999; 99US-0134941.
PR
XX
XX 20-MAY-1999; 99US-0135124.
PR
XX
XX 21-MAY-1999; 99US-0135353.
PR
XX
XX 24-MAY-1999; 99US-0135629.
PR
XX
XX 25-MAY-1999; 99US-0136021.
PR
XX
XX 27-MAY-1999; 99US-0136392.
PR
XX
XX 28-MAY-1999; 99US-0136782.
PR
XX
XX 01-JUN-1999; 99US-0137222.
PR
XX
XX 03-JUN-1999; 99US-0137528.
PR
XX
XX 04-JUN-1999; 99US-0137502.
PR
XX
XX 07-JUN-1999; 99US-0137724.
PR
XX
XX 08-JUN-1999; 99US-0138094.
PR
XX
XX 10-JUN-1999; 99US-0138540.

```
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140354.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147392.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.

PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161982.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 40.0%; Score 6; DB 21; Length 513;
Best Local Similarity 100.0%; Pred. No.1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Caps 0;
QY 5 PHRSTP 10
|||||
```

Db 40 PHRSTP 45

RESULT 133
AAM24230

ID AAM24230 standard; Protein: 525 AA.

XX AAM24230;

DT 17-MAR-1998 (first entry)

DE Human melanoma associated delayed early response variant protein.

XX Melanoma associated delayed early response gene; MADER gene;

KM MADER protein; growth alteration; malignant melanoma; breast carcinoma;

XX cancerous condition; MADER translocation event; MADER immunogen;

XX MADER antigen.

XX Homo sapiens.

PF 30-JAN-1997; 97MO-US01586.

PR 30-JAN-1997; 97US-0593563.

PR 30-JAN-1996; 96US-0593563.

PA (MELC-) MELCOR DIAGNOSTICS INC.

PI Johnson JP;

DR WPI; 1997-402557/37.

DR N-PSDB; AAT77840.

PT Monoclonal antibody which binds to the MADER protein - used for

PS detecting cancerous conditions, such as breast carcinoma

XX Disclosure; Fig 3; 69pp; English.

XX The present sequence represents a novel melanoma associated delayed

CC early response (MADER) protein. This 55 kDa nuclear protein is associated

CC with growth alterations in malignant melanomas and other cancerous

CC conditions, and is over-expressed in human malignant melanomas. Several

CC variants of the protein have been identified (AAM24228-31), the present

CC protein being able to bind erg-1 and inhibit its activity. Chromosomal

CC rearrangement of MADER can be detected by hybridising immobilised

CC chromosomal target DNA, that has been rendered single stranded and is

CC obtained from a cell suspected of having undergone a MADER translocation

CC event with a single stranded oligonucleotide probe complementary to a

CC MADER nucleotide sequence. The probe contains a moiety capable of direct

CC or indirect visualisation. Antibodies raised against the MADER protein

CC can be used for detecting a cancerous condition, particularly melanoma

CC malignancies, and especially a breast carcinoma. A composition comprising

CC a MADER immunogen and a pharmaceutically acceptable vehicle can be used

CC to elicit an immune response against a cell which over-expresses a MADER

CC antigen.

SQ Sequence 525 AA;

Query Match 40.0%; Score 6; DB 18; Length 525;

Best Local Similarity 100.0%; Pred. No. 1.5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPES 12

Db 188 RSTPES 193

RESULT 134
ABG04560
ID ABG04560 standard; Protein: 543 AA.

XX ABG04560;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #4551.

KM Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

PN MO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001MO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSE INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR N-PSDB; AAS68747.

PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity -

XX Claim 20; SEQ ID No 34919; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The

CC polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques

CC to restore normal activity of (II) or to treat disease states involving

CC (II). (II) is useful for generating antibodies against it, detecting or

CC quantitating a polypeptide in tissue, as molecular weight markers and as

CC a food supplement. (II) and its binding partners are useful in medical

CC imaging of sites expressing (II). (I) and (II) are useful for treating

CC disorders involving aberrant protein expression or biological activity.

CC The polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations

CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABG00010-ABG30377 represent novel human

CC Note: The sequence data for this patent did not appear in the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pcl_sequences.

SQ Sequence 543 AA;

Query Match 40.0%; Score 6; DB 22; Length 543;

Best Local Similarity 100.0%; Pred. No. 1.6e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHUGPH 6

Db 235 SHUGPH 240

RESULT 135
ABB93594
ID ABB93594 standard; Protein: 614 AA.

```

XX 31-MAY-2002 (first entry)
DT Herbicidally active polypeptide SEQ ID NO 2805.
DE Herbicidally active polypeptide SEQ ID NO 2805.
XX Herbicidal; plant; agriculture; herbicide.
XX Arabidopsis thaliana.
OS WQ200210210-A2.
XX 07-FEB-2002.
PD 28-AUG-2001; 2001MO-EP09892.
PF 28-AUG-2001; 2001MO-EP09892.
XX 28-AUG-2001; 2001MO-EP09892.
PR (FARB ) BAKER AG.
XX Tietjen K, Weidner M;
PI WPI; 2002-269010/31.
XX Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms.
XX Claim 5; SEQ ID NO 2805; 261pp + Sequence Listing; English.
XX The invention relates to identifying target proteins
CC (AB90790-AB904016) for herbicidally active compounds, comprising
CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
CC useful as herbicides.
XX Sequence 614 AA;
SQ
Query Match 40.0%; Score 6; DB 23; Length 614;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 7 RSTPES 12
DB 516 RSTPES 521

```

RESULT 136

```

AGS3019
ID AAG93019 standard; Protein; 739 AA.
XX
AC AAG93019;
XX
DT 26-SEP-2001 (first entry)
XX
DE C glutamyl protein fragment SEQ ID NO: 6773.
XX
KM Corynebacterium; amino acid synthesis; vitamin; saccharide;
KM organic acid synthesis.
XX
OS Corynebacterium glutamicum.
XX
PN EP108790-A2.
XX
PD 20-JUN-2001.
XX
PF 18-DEC-2000; 2000EP-0127688.
XX
PR 16-DEC-1999; 99JP-0377484.
XX

```

```

PR 07-APR-2000; 2000JP-0159162.
PR 03-AUG-2000; 2000JP-0280988.
XX
XX (RYOW ) KYOWA HAKKO KOGYO KK.
XX
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochial K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
XX WPI; 2001-376931/40.
DR N-PSDB; AAH68238.
XX
XX Novel polynucleotides derived from Corynebacterium, for identifying
PT mutation point of a gene, measuring expression of a gene, analysing
PT expression profile or pattern of a gene and identifying homologous gene
PT
XX
XX Claim 17; SEQ ID NO: 6773; 246pp + Sequence Listing; English.
XX
XX The present invention provides a number of nucleotide and protein
CC sequences from the Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of corynebacterium, measuring expression amount and
CC analysing the expression profile or expression pattern of a gene derived
CC from Corynebacterium, and identifying a homologue of a gene derived
CC from corynebacterium. Corynebacterium are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described
CC in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX
SQ Sequence 739 AA;
Query Match 40.0%; Score 6; DB 22; Length 739;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 10 PESRAA 15
DB 537 PESRAA 542

```

RESULT 137

```

ABP27401
ID ABP27401 standard; Protein; 795 AA.
XX
AC ABP27401;
XX
DT 02-JUL-2002 (first entry)
XX
DE Streptococcus polypeptide SEQ ID NO 3978.
XX
XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
XX group A streptococcus; Streptococcus pyogenes; antibacterial;
XX antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
OS Streptococcus agalactiae.
XX
PN WO200234771-A2.
XX
PD 02-MAY-2002.
XX
PF 29-OCT-2001; 2001MO-GB04789.
XX
XX 27-OCT-2000; 2000GB-0026333.
XX
XX 24-NOV-2000; 2000GB-0028727.
XX
XX 07-MAR-2001; 2001GB-0905640.
XX
XX (CHIR-) CHIRON SPA.
XX
XX (GENO-) INST GENOMIC RES.
XX
XX Telford J, Maignani V, Margarit Ros YI, Grandi G, Fraser C;

```

PI Tettelin H;
XX
XX MPI; 2002-352536/38.
DR N-PSDB; ABN68032.
XX
PT New Streptococcus protein for the treatment or prevention of infection
PT or disease caused by Streptococcus bacteria, such as meningitis, and
PT for detecting a compound that binds to the protein -
XX
XX Claim 1; Page 3552; 4525pp; English.
XX
CC The invention relates to a protein (ABP25413-ABP30895) from group B
CC Streptococcus/GBS (Streptococcus agalactiae) or group A Streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins.
XX
SQ Sequence 795 AA;
XX
Query Match 40.0%; Score 6; DB 23; Length 795;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 9 TPESRA 14
DB 350 TPESRA 355
XX
RESULT 138
AAY28935
ID AAY28935 standard; Protein; 821 AA.
XX
AC AAY28935;
XX
DT 27-SEP-1999 (first entry)
XX
DE Platelet-derived growth factor beta receptor.
XX
KW Chimera gene; chromosome translocation; t(5;14)(q33;q32); CEV14 gene;
KW Platelet-derived growth factor beta receptor; PDGFR-beta.
XX
OS Homo sapiens.
XX
PN JP1187885-A.
XX
PD 13-JUL-1999.
XX
PF 26-DEC-1987; 97JP-0361020.
XX
PS 26-DEC-1997; 97JP-0361020.
XX
PR 26-DEC-1997; 97JP-0361020.
XX
PA (MTP) MITSUBISHI YUKA BCL KK.
XX
DR MPI; 1999-451550/38.
XX
N-PSDB; AAX89336.
XX
XX New DNA - and sensitive method for its detection
PS Claim 6; Page 10-13; 13pp; Japanese.
XX
XX The invention describes a new DNA containing a chimera gene specific to

CC the chromosome translocation t(5;14)(q33;q32) formed by fusing the CEV14
CC gene (1-2287 basepairs of the present sequence) with a platelet-derived
CC growth factor beta receptor (PDGFR-beta) gene. A method for the detection
CC of a chimera gene produced by t(5;14)(q33;q32) in a sample by preparing
CC cDNA from the DNA or mRNA in the sample as the template and using an
CC oligonucleotide designed to bind the CEV14 gene region of the above
CC chimera and an oligonucleotide designed to bind the PDGFR gene region as
CC the primers and detecting the amplified product formed by the PCR is also
CC provided. The method can detect a gene specific to the presence of
CC t(5;14)(q33;q32) specifically at a high sensitivity. The present sequence
CC represents the PDGFR-beta sequence.
XX
SQ Sequence 821 AA;
XX
Query Match 40.0%; Score 6; DB 20; Length 821;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHLGPH 6
DB 419 SHLGPH 424
XX
RESULT 139
AAM93437
ID AAM93437 standard; Protein; 853 AA.
XX
AC AAM93437;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide, SEQ ID NO: 3074.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping.
XX
OS Homo sapiens.
XX
PN EP130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-0114089.
XX
PR 08-JUL-1999; 99JP-0194486.
XX
PR 11-JAN-2000; 2000JP-0118774.
XX
PR 02-MAY-2000; 2000JP-0183765.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR MPI; 2001-524255/58.
XX
N-PSDB; AAK94358.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -
XX
XX Claim 8; SEQ ID NO 3074; 1380pp + sequence listing; English.
XX
CC The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesizing the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesized by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is a polypeptide
CC encoded by a full length human cDNA of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
XX

SQ Sequence 853 AA:

Query Match 40.0%; Score 6; DB 22; Length 853;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 TPESRA 14
 |||||
 DB 789 TPESRA 794

RESULT 140
 AAR30169
 ID AAR30169 standard; Protein; 865 AA.

XX AAR30169;
 XX 07-MAY-1993 (first entry)
 XX Marek's Disease Virus glycoprotein B homologue of HSV.
 XX Marek's Disease Virus glycoprotein B homologue of HSV.
 XX Fowlpox virus; FPV; strain NP; MDV; gBh; recombinant virus;
 XX Herpes Simplex Virus.
 XX Marek's Disease Virus.

XX EP520753-A.
 XX 30-DEC-1992.
 XX 24-JUN-1992; 92EP-0305775.
 XX 28-JUN-1991; 91US-0722860.
 XX 10-DEC-1991; 91US-0803633.
 XX (JAPC) NIPPON ZEON KK.
 XX (USDA) US SEC OF AGRIC.
 XX Lee LF, Li Y, Nazerian K, Ogawa R, Yanagida N;
 XX WPI, 1993-001546/01.

XX Recombinant fowl pox virus contg. Marek's disease virus antigen
 PT gene - used to produce cell-free vaccine against Marek's disease
 PT virus
 XX Example 2; Page 15-19; 30pp; English.

XX The MDV gBh of HSV from a BamHI 13 (5.2kb) and K3 (3.6kb) fragment
 CC of MDV GA strain was cloned into pUC18. A 2.8kb BamHI-SalI
 CC subfragment from 13 fragment and a 1.1kb BamHI-EcoRI subfragment
 CC from K3 fragment were ligated with EcoRI, SalI digested pUC18. The
 CC sequence of the putative MDV gBh was determined by sequencing a set
 CC of deletion mutants. The nucleotide and amino acid sequences were
 CC found to be identical with the published sequences of the gBh of
 CC RBIB strain of MDV (Ross et al., J. Gen. Virol., 70:1789-1894, 1988).
 CC A fragment contg. the entire coding region of MDV gBh was inserted
 CC into pN21729R (see AA034774-Q34778) to produce a recombinant FPV/MDVgBh
 CC virus for immunising chickens. See also AA034780-Q34781.

SQ Sequence 865 AA:

Query Match 40.0%; Score 6; DB 14; Length 865;
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRA 14
 |||||
 DB 174 TPESRA 179

RESULT 141
 AAW36051

ID AAW36051 standard; Protein; 1086 AA.

XX AAW36051;

XX 15-JUL-1998 (first entry)

XX Hybrid Marek's disease virus/M. gallisepticum 40 K-C protein.

XX Chimeric; Marek's disease virus; outer membrane protein; fusion protein;
 XX antigen; vaccine; poultry.

XX Chimeric - Marek's disease gammaherpesvirus.
 XX Chimeric - Mycoplasma gallisepticum.

XX Key Location/Qualifiers

XX Region 1..672 /note="derived from Marek's disease virus gB protein"
 XX Region 693..1086 /note="derived from M. gallisepticum antigen"

XX WO9736924-A1.

XX 09-OCT-1997.

XX 28-MAR-1997; 97MO-JP01084.

XX 29-MAR-1996; 96JP-0103548.

XX (JAPC) NIPPON ZEON KK.

XX Saito S, Tsuzaki Y, Yanagida N;

XX WPI, 1997-503046/46.

XX N-PADB; AAT96596.

XX Fusion protein comprising herpes virus outer membrane protein and
 PT antigenic polypeptide - for prevention of infection by Mycoplasma
 PT gallisepticum, especially in poultry

XX Disclosure; Page 22-30; 51pp; Japanese.

XX This sequence represents the chimeric protein 40 K-C which comprises a
 CC fragment of the Marek's disease virus outer membrane protein gB fused
 CC to an antigenic protein from Mycoplasma gallisepticum. The chimeric
 CC protein can be used in recombinant live vaccines for prevention of
 CC infection by Mycoplasma gallisepticum, especially as the outer membrane
 CC protein shows antigenicity in poultry.

SQ Sequence 1086 AA:

Query Match 40.0%; Score 6; DB 18; Length 1086;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRA 14
 |||||
 DB 174 TPESRA 179

RESULT 142
 ABG20753
 ID ABG20753 standard; Protein; 1089 AA.

XX ABG20753;

XX 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #20744.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO2001/5067-A2.
 PN
 XX
 PD 11-OCT-2001.
 XX
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 XX
 PR 31-MAR-2001; 2000US-0540217.
 XX
 PR 23-AUG-2000; 2000US-0649167.
 XX
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI: 2001-639362/73.
 DR N-PSDB; AAS84940.
 XX
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX
 XX
 PS Claim 20; SEQ ID NO 51112; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 CC
 XX
 SQ Sequence 1089 AA;
 Query Match 40.0%; Score 6; DB 22; Length 1089;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 SHLGPH 6
 DB 656 SHLGPH 661
 RESULT 143
 ABB57338
 ID ABB57338 standard; Protein; 1098 AA.
 XX
 AC ABB57338;
 XX
 DT 07-MAR-2002 (first entry)
 XX
 DE Mouse ischaemic condition related protein sequence SEQ ID NO:943.
 XX
 KM Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
 KM vasospastic ischaemia; ischaemic condition; ischaemic disease.
 OS Mus musculus.
 XX
 XX WO20018188-A2.

XX 22-NOV-2001.
 PD
 XX
 PF 18-MAY-2001; 2001WO-JP04192.
 XX
 PR 18-MAY-2000; 2000JP-0145977.
 XX
 XX
 PA (UYNI-) UNIV. NIHON SCHOOL JURIDICAL PERSON.
 XX
 PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
 XX
 DR WPI: 2002-034733/04.
 DR N-PSDB; ABI99825.
 XX
 PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
 PT expression levels of particular genes defined in the specification or
 PT by determining the expression profile of a gene group comprising these
 PT genes
 XX
 XX
 PS Claim 2; Page 2387-2392; 2690pp; English.
 XX
 CC The present invention describes a method for examining ischaemic
 CC conditions, comprising measuring the expression levels of particular
 CC genes (I) in a test sample or determining the expression profile of a
 CC gene group in the sample comprising genes selected from (I). The method
 CC is useful for examining the ischaemic condition (e.g. compressive
 CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
 CC expression levels of particular genes (ABI99202 to ABI99912, encoding
 CC the protein sequences in ABB57020 to ABB57374) or by determining the
 CC expression profile of a gene group comprising these genes. The
 CC expression levels or expression profiles produced by these genes are
 CC used as an indicator when screening for ischaemic condition-improving
 CC drugs or therapeutics for ischaemic diseases. ABI99913 and ABI99914
 CC represent PCR primers for a mouse ischaemic condition related sequence,
 CC which are used in the exemplification of the present invention.
 XX
 SQ Sequence 1098 AA;
 Query Match 40.0%; Score 6; DB 23; Length 1098;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 SHLGPH 6
 DB 655 SHLGPH 660
 RESULT 144
 AAP90646
 ID AAP90646 standard; protein; 1106 AA.
 XX
 AC AAP90646;
 XX
 DT 20-OCT-1989 (first entry)
 XX
 DE Human platelet-derived growth factor receptor.
 XX
 KM Human platelet derived growth factor receptor; agonist
 KM and antagonist drugs; wound healing; prevents atherosclerosis;
 KM cancer; genetic disorders; antibodies.
 XX
 OS Homo sapiens (human).
 XX
 FH Key Location/Qualifiers
 FH Region 45..47
 FT Region 89..91
 FT Region 103..105
 FT Region 215..217
 FT Region 230..232
 FT Region 292..294
 FT Region 307..309
 FT Region 354..356
 FT Region 371..373

CC which possesses a tyrosine kinase activity. This tyrosine kinase
 CC domain is notable in having an insert of approx. 100 amino acids,
 CC as compared with most other receptor tyrosine kinase domains which
 CC are contiguous or have shorter insert sequences. Fragments of this
 CC sequence between 8 and 400 amino acids comprising one or more PDGF
 CC ligand binding region from the extracellular domain may be used to
 CC bind a PDGF ligand.
 XX
 SO Sequence 1106 AA;
 Query Match 40.0%; Score 6; DB 13; Length 1106;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SHLGFH 6
 Db 656 SHLGFH 661
 RESULT 147
 AAR99690
 ID AAR99690 standard; Protein; 1106 AA.
 XX
 AC AAR99690;
 XX
 DT 11-OCT-1996 (first entry)
 XX
 DE Platelet-derived growth factor receptor.
 XX
 KW Platelet-derived growth factor receptor; PDGF-R; peptide dimer;
 KM protein secretion; agonist; antagonist.
 XX
 OS Homo sapiens.
 XX
 PN EP721983-A1.
 XX
 PD 17-JUL-1996.
 XX
 PF 18-JAN-1989; 89EP-0100787.
 XX
 PR 22-JAN-1988; 88US-0146877.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Bell LA, Kindsvogel WR, Siedzielski AZ;
 XX
 DR WPI: 1996-322833/33.
 DR N-PSDB; AAT34552.
 XX
 PT Prodn. of biologically active peptide dimers, esp. platelet-derived
 PT growth factor receptor analogues - useful for systematic designing
 PT of novel (ant)agonists
 XX
 PS Disclosure; Fig 1; 45pp; English.
 XX
 CC The amino acid sequence (AAR99690) of human platelet-derived growth
 CC factor receptor (PDGF-R) was deduced from a cDNA clone (AAT34552)
 CC isolated from a human diploid dermal fibroblast library. The
 CC cDNA can be used in novel constructs that allow the prodn.
 CC of secreted biologically active PDGF-R analogues. This may
 CC comprise linking a sequence coding for PDGF-R, or the ligand-
 CC binding domain of the PDGF-R extracellular region, to a protein
 CC secretion signal (e.g. SUC2) and promoter, and expression in
 CC transformed host cells, esp. Saccharomyces cerevisiae. The
 CC secreted PDGF-R analogues are used in ligand screening procedures,
 CC to screen for (ant)agonists, and in diagnostic assays.
 XX
 SO Sequence 1106 AA;
 Query Match 40.0%; Score 6; DB 17; Length 1106;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGFH 6
 Db 656 SHLGFH 661
 RESULT 148
 ABG03949
 ID ABG03949 standard; Protein; 1225 AA.
 XX
 AC ABG03949;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #3940.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KM food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSEQ) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI: 2001-639362/73.
 DR N-PSDB; AAS68136.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX
 PS Claim 20; SEQ ID No 34308; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pat_sequences.
 XX
 SO Sequence 1225 AA;
 Query Match 40.0%; Score 6; DB 22; Length 1225;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 8 STEPSR 13
 I|||||

Db 213 STPEER 218

RESULT 149

AA06300 standard; Protein; 2034 AA.

AA06300; 23-AUG-1999 (first entry)

Human activated calcium channel alpha 1H-2 subunit.

Calcium channel subunit 1H-2; human; neurological disorder; endocrinological disorder; cardiovascular disorder; neurological disorder; hepatic disorder; respiratory disorder; vascular disorder.

Homo sapiens.

MO0928342-A2.

10-JUN-1999.

03-DEC-1998; 98MO-US25671.

10-NOV-1998; 98US-0188932.

03-DEC-1997; 97US-0984709.

(SIBI-) SIBIA NEUROSCIENCES INC.

Hans M, Harpold M, Stauderman K, Urrutia A, Washburn MS;

WPI; 1999-371096/31.

N-PSDB; AAX59082.

Subunits of calcium channels

Example 2; Page 162-171; 171pp; English.

The present sequence represents the alpha 1H-2 subunit of a human low-voltage activated calcium channel. Alpha 1H-2 subunit cDNA (see AAX59082) was obtained by PCR amplification from a Tt cell cDNA library. The invention provides calcium channel subunits, including 2 splice variants of alpha 1H-1 (see AAY06298 and AAY06299), and isoform alpha 1H-2. Relative to alpha 1H-1, the 1H-2 isoform contains a 319 amino acid deletion within the intracellular loop between domains II and III. The splice variant deletion was identified by PCR in all cells and tissues examined. These included Tt cells, amygdala, caudate nucleus, putamen, heart, kidney and liver cells. Cells and vectors containing nucleic acid encoding these calcium channel subunits, and methods for identifying compounds that modulate the activity of calcium channels that contain these subunits are provided by the invention. Such compounds are used to treat calcium channel mediated disorders, including neurological, endocrinological, cardiovascular, CC, urological, hepatic, respiratory and vascular disorders.

Sequence 2034 AA;

Query Match 40.0%; Score 6; DB 20; Length 2034;

Best Local Similarity 100.0%; Pred. No. 4.9e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 TPESRA 14

Db 1970 TPESRA 1975

RESULT 150

ABG15478 standard; Protein; 2129 AA.

ABG15478;

18-FEB-2002 (first entry)

Novel human diagnostic protein #15469.

Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.

Homo sapiens.

WO200175067-A2.

11-OCT-2001.

30-MAR-2001; 2001MO-US08631.

31-MAR-2000; 2000US-0540217.

23-AUG-2000; 2000US-0649167.

(HYSE-) HYSEQ INC.

Dremanac RT, Liu C, Tang YT;

WPI; 2001-639362/73.

N-PSDB; AAS79665.

New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity

Claim 20; SEQ ID NO 45837; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations and responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic amino acid sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 2129 AA;

Query Match 40.0%; Score 6; DB 22; Length 2129;

Best Local Similarity 100.0%; Pred. No. 5.1e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLCPH 6

Db 568 SHLCPH 573

RESULT 151

ABG20749 standard; Protein; 2129 AA.

ABG20749;

XX 18-FEB-2002 (first entry)
 XX Novel human diagnostic protein #20740.
 DE Human: chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder.
 KW Homo sapiens.
 OS WO200175067-A2.
 XX PN 11-OCT-2001.
 XX PD 30-MAR-2001; 2001WO-US08631.
 PF 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 PA Drmanac RT, Liu C, Tang YT;
 PI WPI: 2001-639362/73.
 XX DR N-PSDB; AAS84936.
 DR New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX Claim 20; SEQ ID NO 51108; 103pp; English.
 PS The invention relates to isolated polynucleotide (I) and
 XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC XX
 SQ Sequence 2129 AA:
 Query Match 40.0%; Score 6; DB 22; Length 2129;
 Best Local Similarity 100.0%; Pred. No. 5.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 SHLGP 6
 DB 568 SHLGP 573
 RESULT 152
 ID ABB62393 standard; Protein; 2139 AA.
 XX ABB62393;
 AC ABB62393;
 XX 26-MAR-2002 (first entry)
 FT

XX Drosophila melanogaster polypeptide SEQ ID NO 13971.
 DE Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical.
 KW Drosophila melanogaster.
 OS WO200171042-A2.
 XX PN 27-SEP-2001.
 XX PD 23-MAR-2001; 2001WO-US09231.
 PF 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX (PEKE) PE CORP NY.
 PA Venter JC, Adams M, Li PMD, Myers EW;
 PI WPI: 2001-656860/75.
 XX DR N-PSDB; ABL06496.
 DR New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX Disclosure; SEQ ID NO 13971; 21pp + Sequence Listing; English.
 PS The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL01840-ABL16175) and expressed DNA
 CC sequences (AB57737-AB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC XX
 SQ Sequence 2139 AA:
 Query Match 40.0%; Score 6; DB 22; Length 2139;
 Best Local Similarity 100.0%; Pred. No. 5.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 7 RSTPES 12
 DB 1357 RSTPES 1362
 RESULT 153
 ID AAY06298 standard; Protein; 2353 AA.
 XX AAY06298;
 AC AAY06298;
 XX 23-AUG-1999 (first entry)
 DE Human activated calcium channel alpha 1H-1 subunit.
 XX Calcium channel subunit 1H-1; human; neurological disorder;
 KW endocrinological disorder; cardiovascular disorder;
 KW urological disorder; hepatic disorder; respiratory disorder;
 KW vascular disorder.
 XX Homo sapiens.
 OS Homo sapiens.
 FT Key Location/Qualifiers
 FT Region 420..794
 FT /note="intracellular loop"

FT Region 520..528
 FT /note="poly-His region"
 XX
 PN MO928342-A2.
 PD 10-JUN-1999.
 XX
 PF 03-DEC-1998; 98WO-US25671.
 XX
 PR 10-NOV-1998; 98US-0188932.
 PR 03-DEC-1997; 97US-0984709.
 XX
 PA (SIBI-) SIBIA NEUROSCIENCES INC.
 XX
 PI Hans M, Harpold M, Stauderman K, Urrutia A, Washburn MS;
 PI Williams M;
 XX
 DR WPI: 1999-371096/31.
 DR N-PSDB; AAX59080.
 XX
 PT Subunits of calcium channels
 XX
 PS Example 1; Page 138-149; 171pp; English.
 XX
 CC The present sequence represents the the alpha 1H-1 subunit of a human
 CC low-voltage activated calcium channel. The sequence was deduced
 CC from cDNA (see AAX59080) isolated from a TT cell cDNA library.
 CC The invention provides calcium channel subunits, including 2 splice
 CC variants of alpha 1H-1 (see also AAY06299), and isoform alpha 1H-2
 CC (see AAY06300). The 1H-1 splice variants were detected by RT-PCR
 CC in multiple tissues. They differ only at amino acid 2230, being
 CC either Asp or Glu. 1H-2 has a 319-amino acid deletion in the
 CC I-II intracellular loop relative to 1H-1. Cells and vectors
 CC containing nucleic acids encoding these subunits, and methods for
 CC identifying compounds that modulate the activity of calcium channels
 CC that contain the subunits are provided by the invention. Such
 CC compounds are used to treat calcium channel mediated disorders,
 CC including neurological, endocrinological, cardiovascular,
 CC urological, hepatic, respiratory and vascular disorders.
 XX
 SO Sequence 2353 AA:
 Query Match 40.0%; Score 6; DB 20; Length 2353;
 Best Local Similarity 100.0%; Pred. No. 5.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 9 TPESRA 14
 DB 2289 TPESRA 2294

RESULT 154
 AAY06299
 ID AAY06299 standard; Protein; 2353 AA.
 XX
 AC AAY06299;
 XX
 DT 23-AUG-1999 (first entry)
 XX
 DE Human activated calcium channel alpha 1H-1 subunit.
 XX
 DE Human activated calcium channel alpha 1H-1 subunit.
 XX
 KW Calciun channel subunit 1H-1; human; neurological disorder;
 KW endocrinological disorder; cardiovascular disorder;
 KW urological disorder; hepatic disorder; respiratory disorder;
 KW vascular disorder.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FH 420..794
 FT Region /note="intracellular loop"
 FT 520..528
 FT Region /note="poly-His region"

XX
 PN MO928342-A2.
 PD 10-JUN-1999.
 XX
 PF 03-DEC-1998; 98WO-US25671.
 XX
 PR 10-NOV-1998; 98US-0188932.
 PR 03-DEC-1997; 97US-0984709.
 XX
 PA (SIBI-) SIBIA NEUROSCIENCES INC.
 XX
 PI Hans M, Harpold M, Stauderman K, Urrutia A, Washburn MS;
 PI Williams M;
 XX
 DR WPI: 1999-371096/31.
 DR N-PSDB; AAX59081.
 XX
 PT Subunits of calcium channels
 XX
 PS Example 1; Page 151-161; 171pp; English.
 XX
 CC The present sequence represents the the alpha 1H-1 subunit of a human
 CC low-voltage activated calcium channel. The sequence was deduced
 CC from cDNA (see AAX59080) isolated from a TT cell cDNA library.
 CC The invention provides calcium channel subunits, including 2 splice
 CC variants of alpha 1H-1 (see also AAY06298), and isoform alpha 1H-2
 CC (see AAY06300). The 1H-1 splice variants were detected by RT-PCR
 CC in multiple tissues. They differ only at amino acid 2230, being
 CC either Asp or Glu. 1H-2 has a 319-amino acid deletion in the
 CC I-II intracellular loop relative to 1H-1. Cells and vectors
 CC containing nucleic acids encoding these subunits, and methods for
 CC identifying compounds that modulate the activity of calcium channels
 CC that contain the subunits are provided by the invention. Such
 CC compounds are used to treat calcium channel mediated disorders,
 CC including neurological, endocrinological, cardiovascular,
 CC urological, hepatic, respiratory and vascular disorders.
 XX
 SO Sequence 2353 AA:
 Query Match 40.0%; Score 6; DB 20; Length 2353;
 Best Local Similarity 100.0%; Pred. No. 5.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 9 TPESRA 14
 DB 2289 TPESRA 2294

RESULT 155
 ABG30840
 ID ABG30840 standard; Protein; 2353 AA.
 XX
 AC ABG30840;
 XX
 DT 07-OCT-2002 (first entry)
 XX
 DE Human voltage-dependent T-type calcium channel alpha_H subunit protein.
 XX
 DE Human; calcium channel protein; prostate cancer; Parkinson's disease;
 KW CNS; central nervous system disorder; stroke; cardiovascular disorder;
 KW hypertension; angina; haematological disorder; thrombosis; pain relief;
 KW voltage-dependent T-type calcium channel alpha_H subunit; asthma;
 KW inflammatory disease; allergy.
 XX
 OS Homo sapiens.
 XX
 PN MO200252003-A2.
 XX
 PD 04-JUL-2002.
 XX
 PF 19-DEC-2001; 2001WO-EPI5088.
 XX

PR 26-DEC-2000; 2000US-257297P.
 PR 02-APR-2001; 2001US-280113P.
 XX (FARB) BAYER AG.
 XX
 PI Smolyar A;
 XX
 DR WPI: 2002-575380/61.
 XX
 PT Polynucleotides encoding a novel calcium channel protein polypeptide,
 PT useful in preventing or treating diseases related to calcium channel
 PT protein dysfunction, e.g. cancer or CNS, cardiovascular or
 PT haematological disorders -
 XX
 XX Example 5; Fig 3; 144pp; English.
 PS
 CC The present invention relates to a new polynucleotide which encodes a
 CC human calcium channel protein polypeptide. The polypeptide,
 CC polynucleotide and methods are useful in preventing or treating diseases
 CC related to calcium channel protein dysfunction, such as cancers (e.g.
 CC prostate), CNS (central nervous system) disorders (e.g. Parkinson's
 CC disease, stroke), cardiovascular disorders (e.g. hypertension, angina),
 CC or haematological disorders (e.g. thrombosis). The expression vector and
 CC reagent are useful in preparing a medicament for modulating the activity
 CC of the calcium channel protein in a disease. The methods are also useful
 CC in detecting or producing polypeptides and polynucleotides used in
 CC treating the above diseases, and for screening for agents that regulate
 CC the activity of the calcium channel protein. These are also useful in
 CC treating inflammatory diseases (e.g. asthma, allergy) and in relieving
 CC pain from the above diseases. The present amino acid sequence represents
 CC the human voltage-dependent T-type calcium channel alpha_H subunit
 CC protein, as described in the invention.
 XX
 SQ Sequence 2353 AA;
 XX
 Query Match 40.0%; Score 6; DB 23; Length 2353;
 Best Local Similarity 100.0%; Pred. No. 5,6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 9 TPESRA 14
 Db 2289 TPESRA 2294
 XX
 RESULT 156
 AAM22611
 ID AAM22611 standard; Protein: 4545 AA.
 AC AAM22611;
 XX
 DT 02-MAR-1998 (first entry)
 DE Hybrid srmG/tylG ORF1 protein.
 XX
 KW Tyactone synthase gene cluster; tylG gene; multifunctional protein;
 KW polyketide; tyactone synthesis; antibiotic; tylosin; hybrid protein.
 XX
 OS Streptomyces ambofaciens.
 OS Streptomyces fradiae.
 XX
 PH Key
 FT Modified-site 1 Location/Qualifiers
 FT 1
 FT /note="encoded by GTG"
 FT 15..418
 FT /note="ketosynthase domain, KS'(s), from tylG"
 FT 525..882
 FT /note="acyltransferase domain, AT(s), from tylG"
 FT 942..1025
 FT /note="acyl carrier protein domain, ACP(s), from tylG"
 FT 1060..1483
 FT /note="ketosynthase domain, KSL, from srmG"
 FT 1596..1953
 FT /note="acyltransferase domain, AT1, from srmG"

FT Domain
 FT 2232..2416
 FT /note="ketoreductase domain, KR1, from srmG"
 FT 2533..2616
 FT /note="acyl carrier protein domain, ACP1, from srmG"
 FT 2641..3064
 FT /note="ketosynthase domain, KS2, from srmG"
 FT 3184..3520
 FT /note="acyltransferase domain, AT2, from srmG"
 FT 3546..3727
 FT /note="dehydratase domain, DH2, from srmG"
 FT 4083..4268
 FT /note="ketoreductase domain, KR2, from srmG"
 FT 4374..4457
 FT /note="acyl carrier protein domain, ACP2, from srmG"
 FT Domain
 XX
 XX EP791655-A2.
 XX
 XX 27-AUG-1997.
 XX
 XX 19-FEB-1997; 97EP-0301056.
 XX
 XX 22-FEB-1996; 96US-0012078.
 XX
 XX (ELIL) LILLY & CO ELI.
 XX
 XX Dehoff BS, Kuhstoss SA, Rostock PR, Sutton KL;
 XX WPI: 1997-418046/39.
 XX DR N-PSDB; AAT80415.
 XX
 PT DNA encoding Streptomyces fradiae tyactone synthase domain - for
 PT production of tylosin-related polyketide compounds
 XX
 PS Claim 23; Pages 198-212; 220pp; English.
 XX
 CC This sequence represents a hybrid protein of the invention. This
 CC sequence was created by replacing a EcoRI-ApaI fragment of srmG ORF1 with
 CC a EcoRI-SstI fragment from tylG ORF1. The position of the residues from
 CC each of the two genes is not given in the specification. The srmG gene
 CC (see AAT80414) was isolated from Streptomyces ambofaciens, and encodes
 CC the multi-functional proteins which direct the synthesis of the
 CC polyketide platenolide. Platenolide is the basic building block of the
 CC macroide antibiotic spiramycin. The tylG gene (see AAT80413) is the
 CC tyactone synthase gene cluster of the invention. The tylG sequence was
 CC isolated from Streptomyces fradiae, and encodes multifunctional proteins
 CC which direct the synthesis of the polyketide tyactone. Tyactone is the
 CC basic building block of the antibiotic tylosin. The hybrid sequence can
 CC be used to transform S. ambofaciens lacking the srmG ORF1 sequence, or S.
 CC fradiae lacking the tylG ORF1 sequence, so that they can produce
 CC polyketides. The DNA sequence can be modified so as to alter the type of
 CC carboxylic acids incorporated, the number of carboxylic acids
 CC incorporated and/or the post-condensation reactions performed, thereby
 CC resulting in novel tylosin-related polyketides.
 XX
 SQ Sequence 4545 AA;
 XX
 Query Match 40.0%; Score 6; DB 18; Length 4545;
 Best Local Similarity 100.0%; Pred. No. 9,9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 10 PESRAA 15
 Db 4008 PESRAA 4013
 XX
 RESULT 157
 AAM23716
 ID AAM23716 standard; Protein: 4550 AA.
 AC AAM23716;
 XX
 DT 27-FEB-1998 (first entry)

DE Platenolide synthase ORF1 protein.
 XX
 KM Platenolide synthase gene cluster; platenolide production; strmg gene;
 XX multi-functional protein; macrolide antibiotic; spiramycin.
 XX
 OS Streptomyces ambofaciens.
 XX
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1
 FT /note="encoded by GTC"
 FT 15..418
 FT Domain
 FT /note="ketosynthase domain, KS'(s)"
 FT 525..882
 FT Domain
 FT /note="acyltransferase domain, AT(s)"
 FT 942..1025
 FT Domain
 FT /note="acyl carrier protein domain, ACP(s)"
 FT 1060..1483
 FT Domain
 FT /note="ketosynthase domain, KS1"
 FT 1596..1953
 FT Domain
 FT /note="acyltransferase domain, AT1"
 FT 2232..2416
 FT Domain
 FT /note="ketoreductase domain, KR1"
 FT 2533..2616
 FT Domain
 FT /note="acyl carrier protein domain, ACP1"
 FT 2641..3064
 FT Domain
 FT /note="ketosynthase domain, KS2"
 FT 3184..3520
 FT Domain
 FT /note="acyltransferase domain, AT2"
 FT 3546..3727
 FT Domain
 FT /note="dehydratase domain, DH2"
 FT 4083..4268
 FT Domain
 FT /note="ketoreductase domain, KR2"
 FT 4374..4457
 FT Domain
 FT /note="acyl carrier protein domain, ACP2"
 XX
 FT
 FT
 PN EP791656-A2.
 PD 27-AUG-1997.
 XX
 XX
 PF 19-FEB-1997; 97EP-0301066.
 XX
 PR 22-FEB-1996; 96DS-0012050.
 XX
 PA (ELIL) LILLY & CO ELI.
 XX
 PI Burgett SG, Kuhstoss SA, Rao RN, Richardson MA;
 PI Rosteck PR;
 DR WPI: 1997-418047/39.
 DR N-PSDB; AAT78508.
 XX
 PT DNA encoding Streptomyces ambofaciens platenolide synthase domain -
 PT for production of spiramycin-related polyketide antibiotics
 XX
 XX
 PS C1a1n 8; Pages 33-47; 81pp; English.
 XX
 CC AAW23716-W23720 represent proteins encoded by the platenolide synthase
 CC gene cluster of the invention. The gene cluster is also referred to as
 CC the strmg gene, and was isolated from Streptomyces ambofaciens. These
 CC sequences are multi-functional proteins which direct the synthesis of
 CC the polyketide platenolide. Platenolide is the basic building block of
 CC the macrolide antibiotic spiramycin. The DNA can be used to produce
 CC compounds exhibiting antibiotic activity based on the platenolide
 CC structure, including specifically the macrolide antibiotic spiramycin and
 CC spiramycin analogues and derivatives. Modifications of the platenolide
 CC synthase DNA sequence can be made so as to change the number and type of
 CC carboxylic acids incorporated into the growing polyketide chain and to
 CC change the kind of post-condensation processing that is conducted.
 XX
 SQ Sequence 4550 AA;

Query Match 40.0%; Score 6; DB 18; Length 4550;
 Best local Similarity 100.0%; Pred. No. 9.9e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 10 PESRAA 15
 |||||
 Db 4013 PESRAA 4018
 RESULT 158
 AAW22606
 ID AAW22606 standard; Protein; 4550 AA.
 XX
 AC AAW22606;
 XX
 DT 27-FEB-1998 (first entry)
 XX
 DE Platenolide synthase ORF1 protein.
 XX
 XX
 KM Tyllactone synthase gene cluster; tylg gene; multifunctional protein;
 KM polyketide; tyllactone synthesis; antibiotic; tyllactone.
 XX
 OS Streptomyces ambofaciens.
 XX
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1
 FT /note="encoded by GTC"
 FT 15..418
 FT Domain
 FT /note="ketosynthase domain, KS'(s)"
 FT 525..882
 FT Domain
 FT /note="acyltransferase domain, AT(s)"
 FT 942..1025
 FT Domain
 FT /note="acyl carrier protein domain, ACP(s)"
 FT 1060..1483
 FT Domain
 FT /note="ketosynthase domain, KS1"
 FT 1596..1953
 FT Domain
 FT /note="acyltransferase domain, AT1"
 FT 2232..2416
 FT Domain
 FT /note="ketoreductase domain, KR1"
 FT 2533..2616
 FT Domain
 FT /note="acyl carrier protein domain, ACP1"
 FT 2641..3064
 FT Domain
 FT /note="ketosynthase domain, KS2"
 FT 3184..3520
 FT Domain
 FT /note="acyltransferase domain, AT2"
 FT 3546..3727
 FT Domain
 FT /note="dehydratase domain, DH2"
 FT 4083..4268
 FT Domain
 FT /note="ketoreductase domain, KR2"
 FT 4374..4457
 FT Domain
 FT /note="acyl carrier protein domain, ACP2"
 XX
 PN EP791655-A2.
 PD 27-AUG-1997.
 XX
 XX
 PF 19-FEB-1997; 97EP-0301056.
 XX
 PR 22-FEB-1996; 96DS-0012078.
 XX
 PA (ELIL) LILLY & CO ELI.
 XX
 PI Dehoff BS, Kuhstoss SA, Rosteck PR, Sutton KL;
 DR WPI: 1997-418046/39.
 DR N-PSDB; AAT80414.
 XX
 PT DNA encoding Streptomyces fradiae tyllactone synthase domain - for
 PT production of tyllactone-related polyketide compounds
 XX
 XX
 PS Example 2; Pages 135-149; 220pp; English.
 XX
 CC AAW22606-W22610 represent proteins encoded by the platenolide synthase
 CC gene cluster. The gene cluster is also referred to as the strmg gene, and
 CC was isolated from Streptomyces ambofaciens. These sequences are

Query Match	Best Local Similarity	Score 6;	DB 18;	Length 4550;
Matches 6;	Conservative	0;	Mismatches 0;	Indels 0;
Gaps	0;			
QY	10 PESRAA 15			
Db	4013 PESRAA 4018			
RESULT 159				
ID	AAB47749			
	AAB47749 standard; peptide; 8 AA.			
XX	AAB47749:			
AC	21-FEB-2002 (first entry)			
DT	SIV Tat(28-35).			
XX				
DE	Epitope: simian immunodeficiency virus; SIV; cytotoxic T lymphocyte;			
XX	CTL; HIV; viral escape variant; acute phase; vaccine.			
KW	Simian immunodeficiency virus.			
XX				
OS	Simian immunodeficiency virus.			
XX	MO200178775-A2.			
PN	25-OCT-2001.			
PD	09-APR-2001; 2001WO-US11502.			
XX	12-APR-2000; 2000US-196412P.			
PR	(WISC) WISCONSIN ALUMNI RES FOUND.			
XX				
PA	Watkins DI, Allen TM, O'Connor DH, Mothe BR, Vogel TU;			
PI	WPI; 2002-049178/06.			
XX				
XX				
PT	Identifying CTL-inducing epitope(s) from HIV protein, useful for making			
PT	an HIV vaccine, comprises determining the immune response against the			
PT	epitope in order to select viral escape variants during the first 24			
PT	weeks of infection			
XX				
XX	Disclosure; Page 31; 85pp; English.			
XX				
CC	The sequences given in AAB47748-50 represent epitopes derived from			
CC	simian immunodeficiency virus (SIV) which were used in the method of			
CC	the invention. The method allows identification of at least one			
CC	cytotoxic T lymphocyte (CTL)-inducing epitope from at least one			
CC	HIV-protein, and comprises examining the nucleic acid sequence in the			
CC	first 24 weeks after infection, and confirming that an immune response			
CC	directed against the CTL-inducing epitope is capable of selecting for			
CC	viral escape variants during the acute phase of HIV infection.			
CC	The method is useful for designing a rational, novel vaccine			
CC	approach to HIV. The method is particularly useful for making an HIV			

```

CC vaccine.
XX
SQ Sequence 8 AA;

Query Match 33.3%; Score 5; DB 23; Length 8;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STPES 12
   |||||
Db 1 STPES 5

RESULT 160
AAB47751
ID AAB47751 standard; peptide; 8 AA.
AC AAB47751;
AD
AE
AF
AG
AH
AI
AJ
AK
AL
AM
AN
AO
AP
AQ
AR
AS
AT
AU
AV
AW
AX
AY
AZ
BA
BB
BC
BD
BE
BF
BG
BH
BI
BJ
BK
BL
BM
BN
BO
BP
BQ
BR
BS
BT
BU
BV
BW
BX
BY
BZ
CA
CB
CC
CD
CE
CF
CG
CH
CI
CJ
CK
CL
CM
CN
CO
CP
CQ
CR
CS
CT
CU
CV
CW
CX
CY
CZ
DA
DB
DC
DD
DE
DF
DG
DH
DI
DJ
DK
DL
DM
DN
DO
DP
DQ
DR
DS
DT
DU
DV
DW
DX
DY
DZ
EA
EB
EC
ED
EE
EF
EG
EH
EI
EJ
EK
EL
EM
EN
EO
EP
EQ
ER
ES
ET
EU
EV
EW
EX
EY
EZ
FA
FB
FC
FD
FE
FF
FG
FH
FI
FJ
FK
FL
FM
FN
FO
FP
FQ
FR
FS
FT
FU
FV
FW
FX
FY
FZ
GA
GB
GC
GD
GE
GF
GG
GH
GI
GJ
GK
GL
GM
GN
GO
GP
GQ
GR
GS
GT
GU
GV
GW
GX
GY
GZ
HA
HB
HC
HD
HE
HF
HG
HH
HI
HJ
HK
HL
HM
HN
HO
HP
HQ
HR
HS
HT
HU
HV
HW
HX
HY
HZ
IA
IB
IC
ID
IE
IF
IG
IH
II
IJ
IK
IL
IM
IN
IO
IP
IQ
IR
IS
IT
IU
IV
IW
IX
IY
IZ
JA
JB
JC
JD
JE
JF
JG
JH
JI
JJ
JK
JL
JM
JN
JO
JP
JQ
JR
JS
JT
JU
JV
JW
JX
JY
JZ
KA
KB
KC
KD
KE
KF
KG
KH
KI
KJ
KK
KL
KM
KN
KO
KP
KQ
KR
KS
KT
KU
KV
KW
KX
KY
KZ
LA
LB
LC
LD
LE
LF
LG
LH
LI
LJ
LK
LL
LM
LN
LO
LP
LQ
LR
LS
LT
LU
LV
LW
LX
LY
LZ
MA
MB
MC
MD
ME
MF
MG
MH
MI
MJ
MK
ML
MM
MN
MO
MP
MQ
MR
MS
MT
MU
MV
MW
MX
MY
MZ
NA
NB
NC
ND
NE
NF
NG
NH
NI
NJ
NK
NL
NM
NO
NP
NQ
NR
NS
NT
NU
NV
NW
NX
NY
NZ
OA
OB
OC
OD
OE
OF
OG
OH
OI
OJ
OK
OL
OM
ON
OO
OP
OQ
OR
OS
OT
OU
OV
OW
OX
OY
OZ
PA
PB
PC
PD
PE
PF
PG
PH
PI
PJ
PK
PL
PM
PN
PO
PP
PQ
PR
PS
PT
PU
PV
PW
PX
PY
PZ
QA
QB
QC
QD
QE
QF
QG
QH
QI
QJ
QK
QL
QM
QN
QO
QP
QQ
QR
QS
QT
QU
QV
QW
QX
QY
QZ
RA
RB
RC
RD
RE
RF
RG
RH
RI
RJ
RK
RL
RM
RN
RO
RP
RQ
RR
RS
RT
RU
RV
RW
RX
RY
RZ
SA
SB
SC
SD
SE
SF
SG
SH
SI
SJ
SK
SL
SM
SN
SO
SQ
SR
SS
ST
SU
SV
SW
SX
SY
SZ
TA
TB
TC
TD
TE
TF
TG
TH
TI
TJ
TK
TL
TM
TN
TO
TP
TQ
TR
TS
TT
TU
TV
TW
TX
TY
TZ
UA
UB
UC
UD
UE
UF
UG
UH
UI
UJ
UK
UL
UM
UN
UO
UP
UQ
UR
US
UT
UU
UV
UW
UX
UY
UZ
VA
VB
VC
VD
VE
VF
VG
VH
VI
VJ
VK
VL
VM
VN
VO
VP
VQ
VR
VS
VT
VU
VV
VW
VX
VY
VZ
WA
WB
WC
WD
WE
WF
WG
WH
WI
WJ
WK
WL
WM
WN
WO
WP
WQ
WR
WS
WT
WU
WV
WW
WX
WY
WZ
XA
XB
XC
XD
XE
XF
XG
XH
XI
XJ
XK
XL
XM
XN
XO
XP
XQ
XR
XS
XT
XU
XV
XW
XX
XY
XZ
YA
YB
YC
YD
YE
YF
YG
YH
YI
YJ
YK
YL
YM
YN
YO
YP
YQ
YR
YS
YT
YU
YV
YW
YX
YY
YZ
ZA
ZB
ZC
ZD
ZE
ZF
ZG
ZH
ZI
ZJ
ZK
ZL
ZM
ZN
ZO
ZP
ZQ
ZR
ZS
ZT
ZU
ZV
ZW
ZX
ZY
ZZ

Query Match 33.3%; Score 5; DB 23; Length 8;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STPES 12
   |||||
Db 1 STPES 5

RESULT 161

```

```

AAB47758
ID AAB47758 standard; peptide: 8 AA.
XX
AC AAB47758:
XX
DT 21-FEB-2002 (first entry)
XX
DE CTL epitope 94004 rat.
XX
KM Epitope: simian immunodeficiency virus; SIV; cytotoxic T lymphocyte;
KW CTL; HIV; viral escape variant; acute phase; vaccine.
XX
OS Simian immunodeficiency virus.
XX
PN WO200178775-A2.
XX
PD 25-OCT-2001.
XX
PE 09-APR-2001; 2001WO-US11502.
XX
PR 12-APR-2000; 2000US-196412P.
XX
PA (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI Watkins DI, Allen TM, O'Connor DH, Mothe BR, Vogel TU;
XX
DR WPI; 2002-049178/06.
XX
PT Identifying CTL-inducing epitope(s) from HIV protein, useful for making
PT an HIV vaccine, comprises determining the immune response against the
PT epitope in order to select viral escape variants during the first 24
PT weeks of infection
XX
XX
PS Example 2; Page 66; 85pp; English.
XX
CC The sequences given in AAB47751-60 represent epitopes derived from
CC simian immunodeficiency virus (SIV) which were used in the method of
CC the invention. The method allows identification of at least one
CC cytotoxic T lymphocyte (CTL)-inducing epitope from at least one
CC HIV-protein, and comprises examining the nucleic acid sequence in the
CC first 24 weeks after infection, and confirming that an immune response
CC directed against the CTL-inducing epitope is capable of selecting for
CC viral escape variants during the acute phase of HIV infection.
CC The method is useful for designing a rational, novel vaccine
CC approach to HIV. The method is particularly useful for making an HIV
CC vaccine. This epitope affected rapid viral escape by 4 weeks post
CC infection.
XX
SQ Sequence 8 AA:
XX
Query Match 33.3%; Score 5; DB 23; Length 8;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 8 STPEs 12
| | | | |
Db 1 STPEs 5

RESULT 162
AAG88112
ID AAG88112 standard; Peptide: 10 AA.
XX
AC AAG88112:
XX
DT 11-SEP-2001 (first entry)
XX
DE Saccharomyces cerevisiae peptide, SEQ ID NO: 3061.
XX
KW Saccharomyces cerevisiae; complementary peptide; peptide identification;
KW drug discovery; drug design.
XX
OS Saccharomyces cerevisiae.

```

```

XX
PN WO200142276-A1.
XX
PD 14-JUN-2001.
XX
PE 13-DEC-2000; 2000WO-GB04773.
XX
PR 13-DEC-1999; 99GB-0029471.
XX
PA (PROT-) PROTEOM LTD.
XX
PI Roberts GW, Heal JR;
XX
DR WPI; 2001-367863/38.
XX
PT Identifying complementary peptides by analysis of protein and
PT nucleotide sequence databases, useful in drug design -
XX
PS Example 5; Page 452; 488pp; English.
XX
CC The invention relates to the identification of complementary peptides
CC by analysis of protein and nucleotide sequence databases from higher
CC eukaryotic genomes, excluding human and plants. The specific
CC complementary peptides interact with their relevant target proteins
CC encoded in the eukaryote genome. The peptides may be used as reagents
CC and drugs for drug discovery and as lead ligands for drug design and
CC development. The present sequence is a complementary peptide from
CC Saccharomyces cerevisiae.
XX
SQ Sequence 10 AA:
XX
Query Match 33.3%; Score 5; DB 22; Length 10;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 7 STPEs 11
| | | | |
Db 4 STPEs 8

RESULT 163
AAG88113
ID AAG88113 standard; Peptide: 10 AA.
XX
AC AAG88113:
XX
DT 11-SEP-2001 (first entry)
XX
DE Saccharomyces cerevisiae peptide, SEQ ID NO: 3062.
XX
KW Saccharomyces cerevisiae; complementary peptide; peptide identification;
KW drug discovery; drug design.
XX
OS Saccharomyces cerevisiae.
XX
PN WO200142276-A1.
XX
PD 14-JUN-2001.
XX
PE 13-DEC-2000; 2000WO-GB04773.
XX
PR 13-DEC-1999; 99GB-0029471.
XX
PA (PROT-) PROTEOM LTD.
XX
PI Roberts GW, Heal JR;
XX
DR WPI; 2001-367863/38.
XX
PT Identifying complementary peptides by analysis of protein and
PT nucleotide sequence databases, useful in drug design -
XX
PS Example 5; Page 452; 488pp; English.

```

XX The invention relates to the identification of complementary peptides
 CC by analysis of protein and nucleotide sequence databases from higher
 CC eukaryotic genomes, excluding human and plants. The specific
 CC complementary peptides interact with their relevant target proteins
 CC encoded in the eukaryote genome. The peptides may be used as reagents
 CC and drugs for drug discovery and as lead ligands for drug design and
 CC development. The present sequence is a complementary peptide from
 CC *Saccharomyces cerevisiae*.
 XX

SQ Sequence 10 AA;

Query Match 33.3%; Score 5; DB 22; Length 10;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
 |||||
 Db 4 RSTPE 8

RESULT 164

AAR29138
 ID AAR29138 standard; Protein; 11 AA.

XX AAR29138;

DT 19-APR-1993 (first entry)

DE Mutation #10 of T7 polymerase with no exonuclease activity.

KW DNA polymerase; DNA sequencing; base sequencing; chain termination;
 KW Long DNA sequences; ss.

XX Bacteriophage T7.

XX EP516245-A.

PD 02-DEC-1992.

PF 24-DEC-1987; 87EP-0202037.

PR 14-JAN-1987; 87US-0003227.

PR 14-DEC-1987; 87US-0132569.

PA (HARD) HARVARD COLLEGE.

PI Richardson CC, Tabor S;

DR WPI; 1992-400615/49.

DR N-PSDB; AAQ31619.

PT Sequencing method for DNA - using T7 polymerase with no exo
 nuclease activity

PS Example 1; Page 154; 45pp; English.

XX This sequence represents the modified portion (aa214-224) of T7 DNA
 CC polymerase modified by the substitution of His 218 for Ser.
 CC This gives polymerase with little or no associated exonuclease activity,
 CC does not discriminate against nucleotide analogue incorporation and can
 CC utilise small oligonucleotides as specific primers. These are ideal
 CC properties for the manipulation of DNA, partic. for sequencing since
 CC the background level of radioactivity in the PAGE gels is negligible,
 CC there are few or no artificial bands, and the bands are sharp. Such
 CC a polymerase allows novel methods of sequencing long DNA fragments.

SQ Sequence 11 AA;

Query Match 33.3%; Score 5; DB 13; Length 11;
 Best Local Similarity 100.0%; Pred. No. 64;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRAA 15
 |||||
 Db 4 ESRAA 8

RESULT 165

AAR94612
 ID AAR94612 standard; peptide; 11 AA.

XX AAR94612;

DT 21-OCT-1996 (first entry)

DE Protein kinase binding region on c-Fos.

KW Protein kinase; c-Fos; activation; phosphorylation; cell proliferation;
 KW disorder; inhibition.

XX Synthetic.

PN WO9609835-A1.

PD 04-APR-1996.

PF 13-FEB-1995; 95WO-US01770.

PR 29-SEP-1994; 94US-0315067.

PA (REGC) UNIV CALIFORNIA.

PI Deng T, Karin M;

DR WPI; 1996-200718/20.

PT c-Fos regulating protein kinase phosphorylates c-Fos - acts to
 PT potentiates its activity, used to treat a cell proliferation
 PT disorder

PS Claim 23; Page 39; 55pp; English.

XX This sequence comprises the protein binding region on c-Fos which
 CC corresponds to amino acids 226-236 of the mature protein. The
 CC peptide can be useful for the competitive inhibition of naturally
 CC occurring c-Fos in situations where it may be desirable to decrease
 CC the amount of c-Fos activation by phosphorylation by a protein
 CC kinase.

SQ Sequence 11 AA;

Query Match 33.3%; Score 5; DB 17; Length 11;
 Best Local Similarity 100.0%; Pred. No. 64;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPEs 12
 |||||

Db 6 STPEs 10

RESULT 166

AAR34443
 ID AAR34443 standard; peptide; 12 AA.

XX AAR34443;

DT 17-AUG-1993 (first entry)

DE N-terminal sequence of soluble vascular mol. (VCAM).

KW Vascular cell adhesion molecule; VCAM; soluble; plasma.

XX Homo sapiens.

PN WO9307490-A.

XX PD 15-APR-1993.
XX PF 30-SEP-1992; 92WO-GB01791.
XX PR 30-SEP-1991; 91GB-0020767.
XX PA (BRBI-) BRITISH BIO-TECHNOLOGY LTD.
XX PI Gearing AJH;
XX DR WPI: 1993-134626/16.
XX PT Soluble vascular cell adhesion molecule for diagnosis and
PT prognostic(s) - useful in immunoassay as standard for detection
PT indicating vascular damage, e.g. by diabetes or activation of
PT endothelium
XX PS Disclosure; Page 21; 31pp; English.
XX CC Soluble VCAM is naturally present in human serum and plasma.
CC Purification of soluble VCAM from plasma on a monoclonal antibody
CC affinity column reveals three species with mol. wt. of 100kD, 80kD
CC and 40kD. N-terminal sequence analysis of the first 10 residues of
CC the 100kD form (AAR34443) identified the material as being identical
CC to the N-terminus of the published sequence of membrane bound VCAM.
XX SQ Sequence 12 AA;
SQ Query Match 33.3%; Score 5; DB 14; Length 12;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 9 TPESR 13
DB 6 TPESR 10
DB 6 TPESR 10
RESULT 167
AB01192
ID AB01192 standard; Peptide; 12 AA.
XX AC AB01192;
XX DT 13-SEP-2002 (first entry)
DE Human Brn-5 transcriptional regulator peptide fragment.
XX KW Human; cancer; neoplastic disease; tumour specific marker; cytostatic;
KM transcription factor.
XX OS Homo sapiens.
XX PN WO200240716-A2.
XX PD 23-MAY-2002.
XX PF 13-NOV-2001; 2001WO-US43461.
XX PR 16-NOV-2000; 2000US-249508P.
XX PA (CEMI-) CEMINES LLC.
XX PI Palm K;
XX DR WPI: 2002-537346/57.
XX PT Determining the presence of neoplastic molecular markers, by
PT identifying the presence of markers in host test sample using array of
PT neoplastic molecular marker specific reagents and analyzing the array
PT of the reagents
XX PS Example 3; Page 11; 41pp; English.

XX CC The present invention relates to a method for determining the presence of
CC neoplastic molecular markers in a host, involving the use of neoplastic
CC molecular marker specific reagents to detect such markers and analysing
CC the array of reagents, allowing the identification of the neoplastic
CC disease present. This can be used to determine the best treatment for
CC cancers, in particular neural cell, lung and prostate tumours. The
CC present sequence is a peptide derived from a transcription factor capable
CC of acting as a marker of the invention.
XX SQ Sequence 12 AA;
SQ Query Match 33.3%; Score 5; DB 23; Length 12;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 8 STPE 12
DB 3 STPE 7
DB 3 STPE 7
RESULT 168
AAW12798
ID AAW12798 standard; peptide; 13 AA.
XX AC AAW12798;
XX DT 18-APR-1997 (first entry)
DE Antigenic peptide A.
XX KW Antigen; therapy; foot and mouth disease; bovine; cow; horse; globulin.
XX OS Foot and mouth disease virus.
XX PN JP08320324-A.
XX PD 03-DEC-1996.
XX PF 24-MAY-1995; 95JP-0148364.
XX PR 24-MAY-1995; 95JP-0148364.
XX PA (NORQ) NORINSUISANSO KACHIKU EISEI.
XX DR WPI: 1997-073906/07.
XX DT Antigenic peptide for the diagnosis of foot and mouth disease -
DE derived from type 0 foot and mouth disease virus
XX PS Claim 1; Page 5; 7pp; Japanese.
XX CC AAW12798 and AAW12799 represent antigenic peptides derived from type 0
CC foot and mouth disease virus. The reactivity of these peptides was
CC examined by an indirect enzyme antibody method using bovine serum from
CC foot and mouth disease infected cattle. The peptides were combined with
CC horse globulin to give antigens A and B. These sequences can be
CC conjugated to a carrier and used in the diagnosis and treatment of foot
CC and mouth disease. Using these sequences foot and mouth disease can be
CC diagnosed with high sensitivity and specificity.
XX SQ Sequence 13 AA;
SQ Query Match 33.3%; Score 5; DB 18; Length 13;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 RSTPE 11
DB 1 RSTPE 5
DB 1 RSTPE 5
RESULT 169

AA52755
ID AAM52755 standard; peptide; 13 AA.
XX
AC AAM52755;
XX
DT 08-FEB-2002 (first entry)
XX
DE APU G protein-coupled receptor (GPCR) ligand peptide, SEQ ID NO:54.
XX
APU ligand peptide; G protein-coupled receptor; GPCR; arginine rich;
KM Parkinson's disease; Huntington's disease; Pick's disease; circulation;
KM heart; immune system; digestion; metabolism; diabetes; reproduction;
KM AIDS; acquired immunodeficiency syndrome; cardiac; vulnary;
KM immunostimulant; antiviral; neuroprotective; nootropic; antiparkinsonian;
KM anticholinergic; cytosolic; antidiabetic; tranquiliser; antidepressant;
KM anticonvulsant; antiparameic; antihemetic; antitachycardic;
KM antiinflammatory; dermatological; osteopathic; antistimatic;
KM antifertility; arteriosclerotic; hypertensive; hypotensive;
KM cerebroprotective; ophthalmological.
XX
XX Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note="Pyroglutamic acid"
XX
PN WO200170769-A1.
XX
XX 27-SEP-2001.
XX
PE 22-MAR-2001; 2001WO-IP02278.
XX
PR 23-MAR-2000; 2000JP-0087114.
XX 19-SEP-2000; 2000JP-0288691.
XX
PA (TAKE) TAKEDA CHEM IND LTD.
XX
XX Kitada C, Nishizawa N, Hinuma S, Hosoya M;
PI
DR WPI; 2001-648386/74.
XX
PT New peptides rich in arginine and proline are ligands for G
PT protein-coupled receptor protein, useful for treating central nervous
PT system, immune, digestive, metabolic and reproductive disorders,
PT especially human immunodeficiency virus -
XX
XX
PS Claim 16; Page 103; 136pp; Japanese.
XX
CC The invention relates to peptides rich in arginine and proline which
CC are recognised as ligands by the G protein-coupled receptor (GPCR) APU
CC (AAM52755). The peptides of the invention are at least 11 amino acids
CC long, and contain two proline (Pro) residues, two arginine (Arg)
CC residues, a leucine (Leu) residue and a phenylalanine (Phe) residue. They
CC have the general formula:
CC X1-Arg-Pro-Arg-Leu-Phe-X2-X3-Gly-Pro-X4-X5 where:
CC X1 is H or a chain of 1-25 amino acids all optionally substituted in
CC the side chain;
CC X2 is a neutral or basic amino acid residue, both optionally substituted
CC in the side chain;
CC X3 is a neutral, aromatic or basic amino acid residue, all optionally
CC substituted in the side chain;
CC X4 is a peptide bond, a neutral amino acid optionally substituted in the
CC side chain, or an aromatic amino acid; and
CC X5 is an amino acid or a dipeptide (both optionally substituted in the
CC side chain, and whose C terminal carboxy may be reduced to formyl or
CC hydroxymethyl), or OH.
CC The invention also encompasses specific peptides (AAM52707-AAM52722 and
CC AAM52726-AAM52776) conforming to the above formula; esters, amides or
CC salts of peptides of the invention; and pharmaceutical agents containing
CC these peptides or derivatives. The peptides are used as agents to control
CC the function of the central nervous system, circulation, heart, immune

CC system, digestion, metabolism or reproduction. They can also be used for
CC the prevention and treatment of human immunodeficiency virus (HIV)
CC infection and acquired immunodeficiency syndrome (AIDS). The peptides are
CC useful for the prevention and treatment of age-related dementias and
CC Alzheimer's, Parkinson's, Huntington's and Pick's disease; nervous
CC disorders; anxiety or depression; epilepsy; cerebrovascular disorders;
CC acute myocardial infarction; hypercholesterolaemia; arteriosclerosis;
CC hyperlipidaemia; hyperprolactinaemia; hypoglycaemia; diabetes;
CC pancreatitis; secretory, metabolic or toxic disorders (such as thyroid
CC hormone or vitamin B12 deficiency, pituitary deficiency, or alcohol,
CC pharmaceutical or heavy metal poisoning); hyperphagia, polyphagia;
CC infections (such as slow virus diseases and Creutzfeldt-Jacob disease);
CC tumours; kidney disease; trauma; bone fracture; osteoporosis; rheumatoid
CC arthritis; asthma; atopic dermatitis; Turner's syndrome; and infertility.
CC The peptides may also be used as hypertensives and hypotensives, as
CC agents for treating eye pain, and for improving postoperative nutrition.
CC Sequences AAM52707-AAM52722 and AAM52726-AAM52776 represent specifically
CC claimed APU GPCR ligand peptides of the invention.
XX
XX
SQ Sequence 13 AA;
XX

Query Match 33.3%; Score 5; DB 22; Length 13;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGP 5
|||||
Db 6 SHLGP 10

RESULT 170

AA59108
ID AAM59108 standard; peptide; 14 AA.

XX
AC AAM59108;

XX
DT 10-AUG-1998 (first entry)

XX
DE FMDV non-structural viral protein immunogenic peptide fragment #36.

KM Foot and Mouth disease; FMDV; immunogenic; viral nonstructural protein;
KM immunoreactive; antibody; T cell; vaccine; pig; cattle.

XX
OS Sus scrofa.

XX
PN DE19638044-A1.

XX
PD 19-MAR-1998.

XX
PF 18-SEP-1996; 96DE-1038044.

XX
PR 18-SEP-1996; 96DE-1038044.

XX
PA (FARB) BAYER AG.

XX
PI Correa R, Froehlich B, Glatthar-Saalmueller B, Hennen H;
PI Pauy T, Pfaffe E, Saalmueller A, Wiesmueller K;

XX
DR WPI; 1998-180328/17.

XX
PT Vaccines against foot and mouth disease virus - comprising peptide
PT fragments of nonstructural viral proteins

XX
PS Claim 1; Page 22; 27pp; German.

CC AAM59073-W59119 are immunogenic peptides from a foot and mouth disease
CC virus (FMDV) nonstructural protein which are immunoreactive with
CC FMDV-specific antibodies or T cells and can be used in vaccines
CC against the disease. The vaccines are used especially for immunising
CC pigs and cattle.

XX
SQ Sequence 14 AA;

Query Match 33.3%; Score 5; DB 19; Length 14;
 Best Local Similarity 100.0%; Pred. No. 78;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
 |||||
 Db 7 RSTPE 11

RESULT 171

AAW59109
 ID AAW59109 standard; peptide; 14 AA.

XX
 AC AAW59109;

XX
 DT 10-AUG-1998 (first entry)

XX
 DE FMDV non-structural viral protein immunogenic peptide fragment #37.

XX
 KW Foot and Mouth disease; FMDV; immunogenic; viral nonstructural protein;

XX
 KW Immunoreactive; antibody; T cell; vaccine; pig; cattle.

XX
 OS Bos taurus.

XX
 PN DE19638044-A1.

XX
 PD 19-MAR-1998.

XX
 PF 18-SEP-1996; 96DE-1038044.

XX
 PR 18-SEP-1996; 96DE-1038044.

XX
 PA (FARB) BAYER AG.

XX
 PI Correa R, Froehlich B, Glatthaar-Saalmueller B, Heinen H;

XX
 PI Pauly T, Pfaffe E, Saalmueller A, Wiesmueller K;

XX
 DR WPI; 1998-180328/17.

XX
 PT Vaccines against foot and mouth disease virus - comprising peptide

XX
 PS fragments of nonstructural viral proteins

XX
 PS Claim 1; Page 23; 27pp; German.

XX
 CC AAW59073-W59119 are immunogenic peptides from a foot and mouth disease

XX
 CC virus (FMDV) nonstructural protein which are immunoreactive with

XX
 CC FMDV-specific antibodies or T cells and can be used in vaccines

XX
 CC against the disease. The vaccines are used especially for immunising

XX
 CC pigs and cattle.

XX
 SO Sequence 14 AA;

OY 7 RSTPE 11
 |||||
 Db 3 RSTPE 7

RESULT 172

AAAR47671
 ID AAR47671 standard; Protein; 15 AA.

XX
 AC AAR47671;

XX
 DT 27-JUL-1994 (first entry)

XX
 DE HIV epitope #4.

XX
 KW HIV-specific; principal neutralisation; epitope; antigen; conjugate;

XX
 KW outer membrane proteosome; Neisseria; AIDS; vaccine; HIV; infection;

KW disease; antibody; human; monoclonal antibody; 447 antibody.
 XX
 OS Human immunodeficiency virus.

XX
 PN WO9402626-A.

XX
 PD 03-FEB-1994.

XX
 PF 19-JUL-1993; 93WO-US06751.

XX
 PR 20-JUL-1992; 92US-0917212.

XX
 PR 20-JUL-1992; 92US-0917214.

XX
 PR 20-JUL-1992; 92US-0917215.

XX
 PA (MERI) MERCK & CO INC.

XX
 PI Arnold BA, Conley AJ, Keller PM, Shaw AR;

XX
 DR WPI; 1994-048884/06.

XX
 PT New conjugates for treating or preventing HIV infection -

XX
 PT comprising HIV-specific neutralisation epitopes covalently linked

XX
 PS to outer membrane proteosome of Neisseria

XX
 PS Claim 1; Page 93; 181pp; English.

XX
 CC The sequences given in AAR47668-788 are HIV-specific selected

XX
 CC principal neutralisation epitopes which may be used in an antigenic

XX
 CC conjugate linked to purified outer membrane proteosome of Neisseria.

XX
 CC This conjugate may be used in an AIDS vaccine which may be used pre-

XX
 CC and post-exposure to prevent or treat HIV infection or disease. The

XX
 CC vaccine is capable of eliciting specific HIV neutralising antibodies

XX
 CC and bind the broadly neutralising human monoclonal antibody (447

XX
 SO antibody).

XX
 SO Sequence 15 AA;

OY 4 GPHRS 8
 |||||

XX
 Db 11 GPHRS 15

XX
 DT 17-DEC-2001 (first entry)

XX
 DE Tumour suppressor protein 63 peptide fragment.

XX
 KW Tumour suppressor protein 63; cytoskeletal; viral; immunomodulator;

XX
 KW antiinflammatory; haemostatic; gene therapy; tumour; breast carcinoma;

XX
 KW Wilm's tumour; haemopathy; HIV infection; immunological disease;

XX
 KW inflammation; developmental disorder.

XX
 OS Unidentified.

XX
 PN WO200172786-A1.

XX
 PD 04-OCT-2001.

XX
 PF 16-MAR-2001; 2001WO-CN00331.

XX
 PR 17-MAR-2000; 2000CN-0114977.

XX
 PA (BIOW-) BIOWINDOM GENE DEV INC SHANGHAI.

XX Mao Y, Xie Y;
PT WPI; 2001-616469/71.
XX
XX New polypeptide for the diagnosis and treatment of malignant neoplasm,
PT hemopathy, HIV infection, immunological disease and inflammations,
PT comprises the tumor suppressor protein 63 -
XX
XX Example 5; Page 13; 38pp; Chinese.
XX
XX The present invention relates to tumour suppressor protein 63 (see
CC AAG/6900). The tumour suppressor protein and its coding sequence are
CC useful in the diagnosis and treatment of malignant tumours including
CC breast carcinoma and Wilms' tumour, haemopathy, HIV infection,
CC immunological diseases, inflammation and developmental disorders. The
CC present sequence is an N-terminal peptide fragment of the tumour
CC suppressor protein, which was used in an example from the present
CC invention.
XX
XX Sequence 15 AA;
SQ
Query Match 33.3%; Score 5; DB 22; Length 15;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 9 TPESR 13
Db 2 TPESR 6
IIIIII
RESULT 174
ID AAE01628 standard; Protein; 21 AA.
XX
XX AAE01628:
AC
XX
XX 17-JUL-2001 (first entry)
DT
XX
XX Human gene 28 encoded secreted protein HBHME51, SEQ ID NO:178.
DE
XX
XX Human; secreted protein; proliferative disorder; cancer; tumour;
KW foetal abnormality; developmental abnormality; haematopoietic disorder;
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW inflammation; allergy; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
KW cardiovascular disorder; angiogenic disorder; kidney disorder;
KW gastrointestinal disorder; pregnancy-related disorder;
KW endocrine disorder; infection; wound healing; vulnery;
KW cell culture; chemotaxis; food additive; gene therapy;
KW binding partner identification.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Peptide 1..18
FT /label= Signal_peptide
FT 19..21
FT Protein /label= Human_mature_secreted_protein
XX
XX WO200134623-A1.
PN
XX
XX 17-MAY-2001.
PD
XX
XX 01-NOV-2000; 2000WO-US30037.
PF
XX
XX 05-NOV-1999; 9905-0163577.
PR
XX
XX 30-JUN-2000; 2000US-0215137.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Ruben SM, Komatsoulis GA, Moore PA;
PI

XX WPI; 2001-316490/33.
DR N-PDB; AAD05471.
XX
XX Nucleic acids encoding 29 human secreted polypeptides, useful for
PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's
PT disease and diabetic retinopathy -
XX
XX Claim 11; Page 510; 535pp; English.
PS
XX
XX AAD05389-AAD05473 represent cDNAs corresponding to 29 human secreted
CC protein genes, and AAE01546-AAE01630 represent the proteins they encode.
CC AAE01631-AAE01660 represent human secreted protein fragments or variants.
CC The secreted proteins and their genes are useful for preventing,
CC treating or ameliorating medical conditions, e.g., by protein or gene
CC therapy. Pathological conditions can be diagnosed by determining the
CC amount of the new protein in a sample or by determining the presence of
CC mutations in the new genes. Specific uses are described for each of the
CC 29 genes, based on the tissues in which they are most highly expressed,
CC and include developing products for the diagnosis or treatment of
CC proliferative disorders, cancer, tumours, foetal and developmental
CC abnormalities, haematopoietic disorders, diseases of the immune system,
CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
CC allergies, neurological disorders (e.g., Alzheimer's disease,
CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,
CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
CC cardiovascular disorders, angiogenic disorders, kidney disorders,
CC gastrointestinal disorders, pregnancy-related disorders, endocrine
CC disorders, and infections. The proteins can also be used to aid wound
CC healing and epithelial cell proliferation, to prevent skin aging due to
CC sunburn, to maintain organs before transplantation, for supporting cell
CC culture of primary tissues, to regenerate tissues, to identify their
CC cognate ligands or binding partners, and in chemotaxis, and can be used
CC as a food additive or preservative to modify storage properties.
CC Antibodies specific for a protein of the invention can be used in
CC alleviating symptoms associated with the disorders mentioned above, and
CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
CC immunosorbent assay (ELISA). The present sequence represents a human
CC secreted protein of the invention.
XX
XX Sequence 21 AA;
SQ
Query Match 33.3%; Score 5; DB 22; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SHIGP 5
Db 5 SHIGP 9
IIIIII
RESULT 175
ID ABG63804 standard; Protein; 21 AA.
XX
XX ABG63804:
AC
XX
XX 27-AUG-2002 (first entry)
DT
XX
XX Human albumin fusion protein #479.
DE
XX
XX Albumin fusion protein; therapeutic protein X; human albumin; HA;
KW human serum albumin; HSA; cancer; reproductive disorder;
KW digestive disorder; immune disorder; endocrine disorder;
KW haematopoietic disorder; neural disorder; connective disorder;
KW cytostatic; antiinfectivity; antiinflammatory; anticancer;
KW immunomodulator; anti-HIV; antidiabetic; haemostatic; nootropic;
KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
KW osteopathic; antiarthritic.
XX
XX Homo sapiens.
OS
XX
XX Synthetic.
XX

PN WO20017137-A1.
 XX
 PD 18-OCT-2001.
 XX
 PF 12-APR-2001; 2001WO-US11988.
 XX
 PR 12-APR-2000; 2000US-229358P.
 PR 25-APR-2000; 2000US-199384P.
 PR 21-DEC-2000; 2000US-256931P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Haseltine WA;
 XX
 DR WPI; 2002-010886/01.
 PT
 PT New fusion protein for treating disease e.g. diabetes comprises an
 XX albumin fused to a therapeutic protein -
 XX
 PS Claim 1; Page 842; 2102pp; English.
 XX
 CC The present invention relates to albumin fusion proteins comprising a
 CC therapeutic protein X and human albumin (HA, also known as human serum
 CC albumin, HSA). The proteins are useful for treating a disease or
 CC disorder that may be modulated by therapeutic protein X. The albumin
 CC extends the shelf-life of protein X, and may increase its biological
 CC in vitro/in vivo activity. The protein is useful for treating and
 CC diagnosing disorders such as cancer, reproductive disorders, digestive
 CC disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders
 CC (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders
 CC (e.g. diabetes), haematopoietic disorders, neural disorders
 CC (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,
 CC encephalomyelitis, meningitis, schizophrenia), and connective disorders
 CC (e.g. osteoporosis, arthritis). ABG63326-ABG65518 represent albumin
 CC fusion proteins of the invention.
 XX
 SQ Sequence 21 AA;
 XX
 Query Match 33.3%; Score 5; DB 23; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1,1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 SHLGP 5
 II III
 Db 5 SHLGP 9
 XX
 RESULT 176
 ABB39746
 ID ABB39746 standard; Peptide; 34 AA.
 XX
 AC ABB39746;
 XX
 DT 04-FEB-2002 (first entry)
 XX
 DE Peptide #7252 encoded by human foetal liver single exon probe.
 XX
 KM Human; foetal liver; gene expression; single exon nucleic acid probe.
 XX
 OS Homo sapiens.
 OS
 PN WO200157277-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00669.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 XX

PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-483447/52.
 XX
 PR Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human fetal liver -
 XX
 PS Claim 27; SEQ ID NO 32381; 639pp + sequence listing; English.
 XX
 CC The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC fetal liver. The present sequence is a peptide encoded by a single exon
 CC nucleic acid probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 34 AA;
 XX
 Query Match 33.3%; Score 5; DB 22; Length 34;
 Best Local Similarity 100.0%; Pred. No. 1,7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 7 RSTPE 11
 II III
 Db 23 RSTPE 27
 XX
 RESULT 177
 AAM60466
 ID AAM60466 standard; Protein; 34 AA.
 XX
 AC AAM60466;
 XX
 DT 05-NOV-2001 (first entry)
 XX
 DE Human brain expressed single exon probe encoded protein SEQ ID NO: 32571.
 XX
 DE
 XX
 KM Human; brain expressed exon; gene expression analysis; probe;
 KM microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KM epilepsy; cancer.
 XX
 OS Homo sapiens.
 OS
 PN WO200157275-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00667.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-483446/52.
 XX
 PT Single exon nucleic acid probes for analyzing gene expression in human
 XX brains -

PS Example 4; SEQ ID NO: 32571; 650bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX
SQ Sequence 34 AA;

Query Match 33.3%; Score 5; DB 22; Length 34;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPE 11
| | | | |
Db 23 RSTPE 27

RESULT 178
AAM73110
ID AAM73110 standard; Protein; 34 AA.
XX
AC AAM73110;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 33416.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; Leukemia; lymphoma; myeloma.
XX
OS Homo sapiens.
XX
PN MO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00668.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
PS Example 4; SEQ ID NO: 33416; 658bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.
XX
SQ Sequence 34 AA;

Query Match 33.3%; Score 5; DB 22; Length 34;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPE 11
| | | | |
Db 23 RSTPE 27

RESULT 179
AAM33328
ID AAM33328 standard; Protein; 34 AA.
XX
AC AAM33328;
XX
DT 17-OCT-2001 (first entry)
XX
DE Peptide #7365 encoded by probe for measuring placental gene expression.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
XX
OS Homo sapiens.
XX
PN MO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00663.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -
XX
PS Claim 27; SEQ ID NO 33597; 654bp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP:
CC see A131315-A157546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders.
XX
SQ Sequence 34 AA;

Query Match 33.3%; Score 5; DB 22; Length 34;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPE 11
| | | | |
Db 23 RSTPE 27

RESULT 180
ABG42957
ID ABG42957 standard; Peptide; 34 AA.
XX
AC ABG42957;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human peptide encoded by genome-derived single exon probe SEQ ID 32622.

XX Human; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosterosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 OS Homo sapiens.
 PN WO200186003-A2.
 PD 15-NOV-2001.
 PF 30-JAN-2001; 2001WO-US00665.
 PR 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2002-114183/15.
 DR
 XX Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 PS
 XX Claim 27; SEQ ID NO 32622; 634bp; English.

CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression to a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosterosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a peptide/protein

CC encoded by a single exon probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SO Sequence 34 AA;
 QY
 Db 7 RSTPE 11
 11111
 23 RSTPE 27
 Query Match 33.3%; Score 5; DB 23; Length 34;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 181
 ABG22649
 ID ABG22649 standard; Protein; 39 AA.
 XX
 AC ABG22649;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #22640.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 PD 11-OCT-2001.
 PF 30-MAR-2001; 2001WO-US08631.
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 PA (HYSE-) HYSEQ INC.
 PI Drmanac RT, Liu C, Tang YT;
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS6836.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 PS
 XX Claim 20; SEQ ID NO 53008; 103bp; English.

CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (II) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

SO Sequence 39 AA;

Query Match 33.3%; Score 5; DB 22; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRAA 15
|||||

Db 19 ESRAA 23

RESULT 182

ABG27905
ID ABG27905 standard; Protein; 40 AA.

AC ABG27905;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #27896.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR N-PSDB; AAS92092.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

PS Claim 20; SEQ ID No 58264; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.
XX

SO Sequence 40 AA;

Query Match 33.3%; Score 5; DB 22; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GPHRS 8
|||||

Db 17 GPHRS 21

RESULT 183

AAW79463
ID AAW79463 standard; Protein; 47 AA.

AC AAW79463;

DT 04-DEC-1998 (first entry)

DE Staphylococcus aureus protein.

XX Coprea; extensin-like protein; treatment; prevention;
KW bacterial infection; Helicobacter pylori; vaccine.

OS Staphylococcus aureus.

PN WO9823738-A2.

PD 04-JUN-1998.

PF 24-NOV-1997; 97WO-US22092.

PR 25-NOV-1996; 96US-0031469.

PA (SMIK) SMITHKLINE BEECHAM CORP.

PI Warren RL;

DR WPI; 1998-322718/28.

DR N-PSDB; AAW59921.

PT New nucleic acid from Staphylococcus aureus NCIMB 40771 - useful
PT for, e.g. diagnosis, prevention and treatment of bacterial
PT infection(s)

PS Claim 5; Page 103; 114pp; English.

CC AAW79463-67 represent Staphylococcus aureus WCHU (NCIMB 40771)
CC proteins that have homology to a Coprea extensin-like protein
CC fragment. The S. aureus proteins are used to generate antibodies and
CC to screen for antimicrobials. The products are used to treat or prevent
CC bacterial infections, particularly where caused by S. aureus but also
CC against Helicobacter pylori. Particular applications are to treat
CC subjects before surgery or insertion of an in-dwelling device
CC (alternatively the device itself is impregnated before placement).
CC The nucleic acid sequence is used as sources of antisense sequences
CC (for therapeutic use) or regulatory elements for controlling expression
CC of bacterial genes, and for antibacterial screening. The protein can
CC be also used as a vaccine.

SO Sequence 47 AA;

Query Match 33.3%; Score 5; DB 19; Length 47;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 HRSTP 10
|||||

Db 6 HRSTP 10

```
RESULT 184
ABBA0766
ID ABB40766 standard; Peptide: 48 AA.
XX
XX ABB40766;
AC
XX
XX 04-FEB-2002 (first entry)
DT
XX
DE Peptide #8272 encoded by human foetal liver single exon probe.
XX
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS Homo sapiens.
XX
XX WO200157277-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00669.
PE
XX
XX 04-FEB-2000; 2000US-0180312.
PR
XX 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2001-483447/52.
DR
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver -
XX
XX
XX Claim 27; SEQ ID NO 33401; 639pp + sequence listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a peptide encoded by a single exon
CC nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
XX Sequence 48 AA:
SO
Query Match 33.3%; Score 5; DB 22; Length 48;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 4 GPHRS 8
| | | | |
Db 20 GPHRS 24
RESULT 185
AAM61626
ID AAM61626 standard; Protein: 48 AA.
XX
XX AAM61626;
AC
XX
XX 05-NOV-2001 (first entry)
DT
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 33731.
XX
XX Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.
PR
```

```
XX
OS Homo sapiens.
XX
XX WO200157275-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00667.
PE
XX
XX 04-FEB-2000; 2000US-0180312.
PR
XX 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2001-483446/52.
DR
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
XX
XX Example 4; SEQ ID NO: 33731; 650pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX
XX
XX Sequence 48 AA:
SO
Query Match 33.3%; Score 5; DB 22; Length 48;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 4 GPHRS 8
| | | | |
Db 20 GPHRS 24
RESULT 186
AAM74418
ID AAM74418 standard; Protein: 48 AA.
XX
XX AAM74418;
AC
XX
XX 06-NOV-2001 (first entry)
DT
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 34724.
XX
XX Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX
XX Homo sapiens.
OS
XX
XX WO200157276-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00668.
PE
XX
XX 04-FEB-2000; 2000US-0180312.
PR
XX 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR
```

PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488900/53.
DR
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
PS Example 4; SEQ ID NO: 34724; 658bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.
XX
SQ Sequence 48 AA;
XX
Query Match 33.3%; Score 5; DB 22; Length 48;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 4 GPHRS 8
DB 20 GPHRS 24
XX
RESULT 187
AAM34532
ID AAM34532 standard; Protein; 48 AA.
XX
AC AAM34532;
XX
DT 17-OCT-2001 (first entry)
XX
DE Peptide #8569 encoded by probe for measuring placental gene expression.
XX
XX Probe: microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
XX
OS Homo sapiens.
OS
XX WO200157272-A2.
PN
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00663.
PF
XX
XX 04-FEB-2000; 2000US-0180312.
PR
XX 26-MAY-2000; 2000US-0207456.
PR
XX 30-JUN-2000; 2000US-0608408.
PR
XX 03-AUG-2000; 2000US-063366.
PR
XX 21-SEP-2000; 2000US-0234687.
PR
XX 27-SEP-2000; 2000US-0236359.
PR
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2001-48897/53.
DR
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -
XX
PS Claim 27; SEQ ID No 34801; 654bp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP:

CC see AAI31315-AA157546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders.
XX
SQ Sequence 48 AA;
XX
Query Match 33.3%; Score 5; DB 22; Length 48;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 4 GPHRS 8
DB 20 GPHRS 24
XX
RESULT 188
AAB42306
ID AAB42306 standard; Protein; 50 AA.
XX
AC AAB42306;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human ORFX ORF2070 polypeptide sequence SEQ ID NO:4140.
DE
XX
XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnery; antiparinsonian; antiparinsonian; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antineumatic; antihypertensive;
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive.
XX
XX Homo sapiens.
OS
XX WO200058473-A2.
PN
XX 05-OCT-2000.
PD
XX
XX 31-MAR-2000; 2000WO-US08621.
PF
XX
XX 31-MAR-1999; 99US-0127607.
PR
XX 02-APR-1999; 99US-0127636.
PR
XX 05-APR-1999; 99US-0127728.
PR
XX 30-MAR-2000; 2000US-0540763.
XX
XX (CURA-) CURAGEN CORP.
PA
XX Shinkets RA, Leach M;
PI
XX WPI; 2000-602362/57.
DR
XX N-PSDB; AAC76515.
DR
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX
XX Claim 11; Page 3328; 5507bp; English.
PS
XX
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
CC antiparinsonian; antiparinsonian; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC

CC Immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antifungal; antineumatic;
CC antihypoid; and antianemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease, to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.

CC
XX
SQ Sequence 50 AA;

Query Match 33.3%; Score 5; DB 21; Length 50;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRA 14
|||||
Db 20 PESRA 24

RESULT 189
ABP05658
ID ABP05658 standard; Protein: 50 AA.
XX
AC ABP05658;
XX
DT 24-JUN-2002 (first entry)
XX
DE Human ORFX protein sequence SEQ ID NO:11298.
XX
XX Human: open reading frame; ORFX: gene therapy; cancer; cirrhosis;
KM hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
KM degenerative disorder; osteoarthritis; neurodegenerative disorder;
KM cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KM hypertension; hypothyroidism; cholesterol ester storage disease;
KM immune deficiency; immune disorder; infectious disease;
KM autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
KM myasthenia gravis.
XX
OS Homo sapiens.
XX
PN WO200192523-A2.
XX
PD 06-DEC-2001.
XX
PF 29-MAY-2001; 2001WO-US10836.
XX
PR 30-MAY-2000; 2000US-206132P.
PR 29-AUG-2000; 2000US-228716P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach MD;
XX
DR WPI: 2002-106308/14.
DR N-PSDB: ABN21410.
XX
XX Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and autoimmune disorders -
XX
XX Disclosure; SEQ ID 11298; 1037pp; English.
XX
CC The present invention describes substantially purified human proteins
CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1

CC in the specification). ABN15762 to ABN27252 encode the human ORFX
CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
CC treating or preventing a pathology associated with an ORFX-associated
CC disorder in humans, and in the manufacture of a medicament for treating a
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC sequences can be used in gene therapy. ORFX sequences can be used in the
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC osteoarthritis, neurodegenerative disorders, disorders related to organ
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC storage disease, various immune deficiencies and disorders, infectious
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
CC
XX
SQ Sequence 50 AA;

Query Match 33.3%; Score 5; DB 23; Length 50;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HLGPH 6
|||||
Db 34 HLGPH 38

RESULT 190
AAG57801
ID AAG57801 standard; Protein: 54 AA.
XX
AC AAG57801;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 74532.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
XX Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135533.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136392.
PR 01-JUN-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140921.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159299.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.

PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 33.3%; Score 5; DB 21; Length 54;
Best Local Similarity 100.0%; Pred. No. 2.5e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 5; Conservative 0;

QY 9 TPESR 13
DB 13 TPESR 17

RESULT 191

ID ABP10197 standard; Protein; 54 AA.

XX ABP10197;
XX 24-JUN-2002 (first entry)

DE Human ORFX protein sequence SEQ ID NO:20376.

XX Human: open reading frame: ORFX; gene therapy: cancer; cirrhosis;
XX hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
XX degenerative disorder; osteoarthritis; neurodegenerative disorder;
XX cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
XX hypertension; hypothyroidism; cholesterol ester storage disease;
XX immune deficiency; immune disorder; infectious disease;
XX autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
XX myasthenia gravis.

XX Homo sapiens.

XX WO200192523-A2.

XX 06-DEC-2001.

XX 29-MAY-2001; 2001WO-US10836.

XX 30-MAY-2000; 2000US-206132P.

XX 29-AUG-2000; 2000US-228716P.

XX (CURA-) CURAGEN CORP.

XX Shinkets RA, Leach MD;

XX WPI: 2002-106308/14.

XX N-PSDB: ABN25949.

XX Novel human polypeptides and polynucleotides useful for diagnosing,
XX preventing and treating cardiovascular disease, neurodegenerative,
XX hyperproliferative disorders and autoimmune disorders

XX disclosure; SEQ ID 20376; 10377P; English.

XX The present invention describes substantially purified human proteins
XX (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1

CC in the specification). ABN15762 to ABN27252 encode the human ORFX
CC proteins given in ABP00010 to ABP1500. ORFX proteins are useful for
CC treating or preventing a pathology associated with an ORFX-associated
CC disorder in humans, and in the manufacture of a medicament for treating a
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC sequences can be used in gene therapy. ORFX sequences can be used in the
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC osteoarthritis, neurodegenerative disorders, disorders related to organ
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC storage disease, various immune deficiencies and disorders, infectious
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC arthritis, autoimmune inflammatory eye disease. ORFX proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage.

CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 54 AA;

QY 6 HRSTP 10 33.3%; Score 5; DB 23; Length 54;
Best Local Similarity 100.0%; Pred. No. 2.5e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 5; Conservative 0;

DB 30 HRSTP 34

RESULT 192

ID AAU60387 standard; Protein; 55 AA.

XX AAU60387;

XX 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #21283.

XX SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;

XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;

XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;

XX dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US12865.

XX 21-APR-2000; 2000US-199047P.

XX 02-JUN-2000; 2000US-208841P.

XX 07-JUL-2000; 2000US-216747P.

XX (CORI-) CORIXA CORP.

XX Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;

XX L'valsosneuve J, Zhang Y, Jen S, Carter D;

XX WPI: 2001-616774/71.
XX N-PSDB: AAS59609.
XX Propionibacterium acnes polypeptides and nucleic acids useful for
XX treating acne vulgaris -

XX PS Example 1; SEQ ID No 21582; 1069bp; English.
XX CC Sequences AAU9105-AAU68017 represent *Propionibacterium acnes* immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC *P. acnes*. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC *P. acnes* is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of *P. acnes* in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for *P. acnes* proteins. These antibodies can be used to
CC downregulate expression and activity of *P. acnes* polypeptides and
CC therefore treat *P. acnes* infections. The antibodies may also be used as
CC diagnostic agents for determining *P. acnes* presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 55 AA;

Query Match 33.3%; Score 5; DB 22; Length 55;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PRST 9
 |||||
DB 1 PRST 5

RESULT 193
AAM99608
ID AAM99608 standard; Protein; 56 AA.
XX AAM99608;
AC
XX
DT 07-JAN-2002 (first entry)
XX
DE Human excretory related polypeptide SEQ ID NO 345.
XX
KW Human; nocotropic; neuroprotective; cytosolic; dermatological; virucide;
KW immunosuppressive; anti-inflamatory; anti-HIV; antibacterial; vulnary;
KW antiparkinsonian; antiskilling; antianaemic; antirheitic; cancer;
KW antilithematic; hepatotropic; cerebroprotective; antinflammatory;
KW antiallergic; antidiabetic; antidiuretic; anticonvulsant; antitungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine;
KW excretory system.
XX
OS Homo sapiens.
XX
PN WO20015313-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01323.
XX
PR 31-JAN-2000; 2000US-019065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0188874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 03-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235835.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239933.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.

PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246619.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-465569/50.
XX N-PSDB; AA198581.
XX
XX Isolated nucleic acid molecule encoding excretory system antigen is
XX used in preventing, treating or ameliorating a medical condition -
XX
PS Claim 11: SEQ ID NO 345; 574pp + Sequence Listing; English.
XX
XX The invention relates to novel excretory system related human
XX polynucleotides (AA198567-AA199503) and the encoded proteins
XX (AA99594-AA99913) useful for preventing, treating or ameliorating
XX medical conditions e.g. by protein or gene therapy, especially
XX disorders related to the excretory system. The genes are isolated
XX from a range of human tissues disclosed in the specification. The
XX CC nucleic acids, proteins, antibodies and (ant)agonists are useful in
XX the diagnosis, treatment and prevention of: (a) cancer, e.g. breast

CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC hemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 56 AA;
Query Match 33.3%; Score 5; DB 22; Length 56;
Best Local Similarity 100.0%; Pctd. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 11 ESRAA 15
|||
Db 25 ESRAA 29
RESULT 194
ID AAM42423
AA AAM42423 standard; Protein: 56 AA.
XX
AC AAM42423;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human kidney related polypeptide SEQ ID NO 292.
XX
XX Human; kidney antigen; immunosuppressive; antiarthritic; antirheumatic;
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
KW neurotropic; neuroprotective; antibacterial; virucide; fungicide;
KW ophthalmological; antiallergic; hepatotropic; antidiabetic;
KW antiinflammatory; antileuc; vulnerary; anticonvulsant; antiparasitic;
KW gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection.
XX
OS Homo sapiens.
XX
PN WO200155323-A2.
XX
PD 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01343.
PF
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217467.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.

CC Printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 56 AA;
Query Match 33.3%; Score 5; DB 22; Length 56;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 11 ESRRA 15
Db 25 ESRRA 29
RESULT 195
AAI65001
ID AAY65001 standard; Protein: 58 AA.
XX
AC AAY65001;
XX
DT 01-FEB-2000 (first entry)
XX
DE Human 5' EST related polypeptide SEQ ID NO:1162.
XX
KM Human: 5' EST; expressed sequence tag; secreted protein; diagnosis;
KM gene therapy; chromosome mapping; upstream regulatory sequence;
KM forensic; location; development; protein synthesis; stability;
KM regulation; identification.
XX
OS Homo sapiens.
XX
PN WO953051-A2.
XX
PD 21-OCT-1999.
XX
PF 09-APR-1999; 99MO-IB00712.
XX
PR 09-APR-1998; 98US-0057719.
PR 28-APR-1998; 98US-0069047.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI: 2000-038446/03.
DR N-PSDB; AA242615.
XX
PT Novel secreted protein 5' expressed sequence tag sequences used in
PT diagnostic, forensic, gene therapy, and chromosome mapping procedures
XX
PS Claim 3; Page 704; 837pp: English.
XX
AA242265 to AA243075 represent novel 5' expressed sequence tag (EST)
CC sequences, corresponding to human secreted proteins. AAY64651 to
CC AAY65438 represent the EST-related proteins corresponding to AA242265 to
CC AA243052. The 5' ESTs can be used for producing secreted human gene
CC products. They can be used to identify and isolate 5' untranslated
CC regions (UTRs) and upstream regulatory regions which control the
CC location, development stage, rate, and quantity of protein synthesis, as
CC well as stability of mRNA. The ESTs are also useful as probes for
CC chromosome mapping, and to obtain full length cDNA clones. The ESTs can
CC also be used in forensic procedures to identify individuals, or in
CC diagnostic procedures to identify individuals having genetic diseases
CC resulting from abnormal gene expression. The products may also be used in
CC gene therapy protocols. The nucleic acids encoding signal peptides can be
CC used for directing extracellular secretion of a polypeptide or the
CC insertion of a polypeptide into a membrane, or importing a polypeptide
CC into a cell. The proteins encoded by the EST sequences may be useful in
CC treating a variety of human conditions. Secreted proteins have
CC therapeutic value, and the identification of new secreted proteins is
CC valuable. AA242249 to AA242264 and AAY64644 to AAY64650 represent
CC sequences used in the exemplification of the present invention.
XX

SQ Sequence 58 AA;
Query Match 33.3%; Score 5; DB 21; Length 58;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 11 ESRRA 15
Db 32 ESRRA 36
RESULT 196
ABG22141
ID ABG22141 standard; Protein: 58 AA.
XX
AC ABG22141;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #22132.
XX
KM Human: chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PE 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR N-PSDB; AAS86328.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
PS Claim 20; SEQ ID NO 52500; 103pp: English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridization probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 58 AA;

Query Match 33.3%; Score 5; DB 22; Length 58;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 3 LGPHR 7
|||||
Db 32 LGPHR 36

RESULT 197
AAM78786
ID AAM78786 standard; Protein; 58 AA.
XX
AC AAM78786;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 1448.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US04098.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0634936.
PR 15-SEP-2000; 2000US-0663361.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
DR WPI: 2001-476283/51.
DR N-PSDB: AAK51919.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX
PS Claim 20; Page 3715; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAK80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
SQ Sequence 58 AA;

Query Match 33.3%; Score 5; DB 22; Length 58;

Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 4 GPHRS 8
|||||
Db 34 GPHRS 38

RESULT 198
AAU46591
ID AAU46591 standard; Protein; 59 AA.
XX
AC AAU46591;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #7487.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skelley YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI: 2001-616774/71.
DR N-PSDB: AAS59534.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
PS Example 1; SEQ ID NO 7786; 1063pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 59 AA;

Query Match 33.3%; Score 5; DB 22; Length 59;

Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 HRSTP 10
|||||
Db 25 HRSTP 29

RESULT 199

AAU40678
ID AAU40678 standard; Protein: 60 AA.

XX
AC AAU40678;

XX
DT 13-FEB-2002 (first entry)

XX
DE Propionibacterium acnes immunogenic protein #1574.

XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant;

XX
OS Propionibacterium acnes.

XX
PN W0200181581-A2.

XX
PD 01-NOV-2001.

XX
PF 20-APR-2001; 2001WO-US12865.

XX
PR 21-APR-2000; 2000US-199047P.

XX
PR 02-JUN-2000; 2000US-208841P.

XX
PR 07-JUL-2000; 2000US-216747P.

XX
PA (CORI-) CORIXA CORP.

XX
PI Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;

XX
PI L'malsonneuve J, Zhang Y, Jen S, Carter D;

XX
DR WPI; 2001-616774/71.

XX
DR N-PSDB; AAS59513.

XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for

XX
PT vaccinating against and diagnosing infections, especially useful for

XX
PT treating acne vulgaris -

XX
PS Example 1; SEQ ID NO 1873; 1069pp; English.

XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 60 AA;

Query Match 33.3%; Score 5; DB 22; Length 60;

Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 GPHRS 8
|||||
Db 7 GPHRS 11

RESULT 200

AAU64335
ID AAU64335 standard; Protein: 60 AA.

XX
AC AAU64335;

XX
DT 27-FEB-2002 (first entry)

XX
DE Propionibacterium acnes immunogenic protein #25231.

XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.

XX
OS Propionibacterium acnes.

XX
PN W0200181581-A2.

XX
PD 01-NOV-2001.

XX
PF 20-APR-2001; 2001WO-US12865.

XX
PR 21-APR-2000; 2000US-199047P.

XX
PR 02-JUN-2000; 2000US-208841P.

XX
PR 07-JUL-2000; 2000US-216747P.

XX
PA (CORI-) CORIXA CORP.

XX
PI Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;

XX
PI L'malsonneuve J, Zhang Y, Jen S, Carter D;

XX
DR WPI; 2001-616774/71.

XX
DR N-PSDB; AAS59641.

XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for

XX
PT vaccinating against and diagnosing infections, especially useful for

XX
PT treating acne vulgaris -

XX
PS Example 1; SEQ ID NO 25530; 1069pp; English.

XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 60 AA;

Query Match 33.3%; Score 5; DB 22; Length 60;

Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 HRSTP 10
|||||
Db 3 HRSTP 7

Search completed: March 10, 2003, 14:27:43
Job time : 73 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2003, 14:26:25 ; Search time 24 seconds
(without alignments)
18.389 Million cell updates/sec

Title: US-09-689-159a-2_COPY_346_360
Perfect score: 15
Sequence: 1 SHLGPRTSPESRAA 15

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 0

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 200 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15	100.0	15	2	US-08-967-101-170
2	15	100.0	15	2	US-08-592-541-170
3	15	100.0	15	3	US-09-124-698-170
4	15	100.0	15	4	US-09-127-480-170
5	15	100.0	15	4	US-08-496-841C-167
6	15	100.0	15	4	US-09-124-523-170
7	15	100.0	15	4	US-08-875-972-4
8	15	100.0	15	2	US-08-788-231A-15
9	15	100.0	15	3	US-08-670-479-18
10	15	100.0	15	3	US-08-670-964-4
11	15	100.0	15	3	US-08-888-077A-4
12	15	100.0	15	3	US-08-888-077A-4
13	15	100.0	15	2	US-08-967-101-2
14	15	100.0	15	2	US-08-967-101-4
15	15	100.0	15	2	US-08-592-541-2
16	15	100.0	15	2	US-08-592-541-4
17	15	100.0	15	2	US-08-592-541-4
18	15	100.0	15	2	US-08-592-541-134
19	15	100.0	15	3	US-08-923-454A-10
20	15	100.0	15	3	US-08-670-964-2
21	15	100.0	15	3	US-08-888-077A-2
22	15	100.0	15	3	US-08-888-077A-17
23	15	100.0	15	3	US-09-124-698-2
24	15	100.0	15	3	US-09-124-698-4
25	15	100.0	15	3	US-09-124-698-134
26	15	100.0	15	4	US-09-127-480-2
27	15	100.0	15	4	US-09-127-480-4

28	15	100.0	467	4	US-09-127-480-134	Sequence 134, App
29	15	100.0	467	4	US-08-496-841C-2	Sequence 2, App1
30	15	100.0	467	4	US-08-496-841C-4	Sequence 4, App1
31	15	100.0	467	4	US-08-496-841C-134	Sequence 134, App
32	15	100.0	467	4	US-08-496-841C-136	Sequence 136, App
33	15	100.0	467	4	US-08-706-344C-2	Sequence 2, App1
34	15	100.0	467	4	US-08-706-344C-4	Sequence 4, App1
35	15	100.0	467	4	US-08-706-344C-28	Sequence 28, App1
36	15	100.0	467	4	US-08-706-344C-30	Sequence 30, App1
37	15	100.0	467	4	US-08-706-344C-32	Sequence 32, App1
38	15	100.0	467	4	US-08-832-667-3	Sequence 3, App1
39	15	100.0	467	4	US-09-227-725A-1	Sequence 1, App1
40	15	100.0	467	4	US-09-124-523-2	Sequence 2, App1
41	15	100.0	467	4	US-09-124-523-4	Sequence 4, App1
42	15	100.0	467	4	US-09-124-523-134	Sequence 134, App
43	15	100.0	467	4	US-09-375-318-3	Sequence 3, App1
44	15	100.0	256	4	US-09-375-322A-57	Sequence 57, App1
45	15	100.0	270	2	US-07-857-224B-76	Sequence 76, App1
46	15	100.0	359	1	US-08-844-055-2	Sequence 2, App1
47	15	100.0	359	3	US-09-006-849-2	Sequence 2, App1
48	15	100.0	429	1	US-08-906-744A-2	Sequence 2, App1
49	15	100.0	429	3	US-09-093-134-2	Sequence 2, App1
50	15	100.0	566	1	US-08-810-116-11	Sequence 11, App1
51	15	100.0	566	2	US-07-930-548A-11	Sequence 11, App1
52	15	100.0	865	1	US-07-803-533A-13	Sequence 13, App1
53	15	100.0	1106	1	US-08-180-195-2	Sequence 2, App1
54	15	100.0	1106	1	US-08-168-917-2	Sequence 2, App1
55	15	100.0	1106	1	US-08-477-329-2	Sequence 2, App1
56	15	100.0	1106	2	US-08-475-558-2	Sequence 2, App1
57	15	100.0	1106	2	US-08-460-510-2	Sequence 2, App1
58	15	100.0	1106	2	US-08-460-490-2	Sequence 2, App1
59	15	100.0	1106	3	US-08-960-400-2	Sequence 2, App1
60	15	100.0	1106	3	US-08-462-728-4	Sequence 4, App1
61	15	100.0	1106	4	US-09-583-459A-2	Sequence 2, App1
62	15	100.0	1106	4	US-09-583-210-2	Sequence 2, App1
63	15	100.0	1106	4	US-09-583-449A-2	Sequence 2, App1
64	15	100.0	1106	4	US-09-435-059-2	Sequence 2, App1
65	15	100.0	1106	4	US-08-461-917-4	Sequence 4, App1
66	15	100.0	1106	5	PCT-US92-00730-2	Sequence 2, App1
67	15	100.0	1106	5	PCT-US92-00862-2	Sequence 2, App1
68	15	100.0	2353	4	US-08-984-709A-50	Sequence 50, App1
69	15	100.0	4545	2	US-08-804-227C-14	Sequence 14, App1
70	15	100.0	4550	2	US-08-804-227C-8	Sequence 8, App1
71	15	100.0	4550	2	US-08-804-198-2	Sequence 2, App1
72	15	100.0	4550	2	US-08-707-873-1	Sequence 1, App1
73	15	100.0	11	2	US-08-707-874-1	Sequence 1, App1
74	15	100.0	11	2	US-08-315-067-1	Sequence 1, App1
75	15	100.0	11	3	US-09-193-797-1	Sequence 1, App1
76	15	100.0	11	5	PCT-US95-01770-1	Sequence 1, App1
77	15	100.0	13	1	US-08-630-897-1	Sequence 1, App1
78	15	100.0	15	5	PCT-US93-06751-4	Sequence 4, App1
79	15	100.0	26	2	US-08-751-767A-67	Sequence 67, App1
80	15	100.0	40	1	US-07-743-518-18	Sequence 18, App1
81	15	100.0	111	4	US-08-466-886-41	Sequence 41, App1
82	15	100.0	111	4	US-08-469-617-41	Sequence 41, App1
83	15	100.0	161	4	US-09-134-001C-4265	Sequence 4265, App1
84	15	100.0	170	3	US-09-069-896-1	Sequence 1, App1
85	15	100.0	170	4	US-09-471-168-1	Sequence 1, App1
86	15	100.0	242	4	US-09-413-814-98	Sequence 98, App1
87	15	100.0	253	4	US-09-550-058-6	Sequence 6, App1
88	15	100.0	282	3	US-08-911-853-37	Sequence 37, App1
89	15	100.0	282	4	US-09-479-409-37	Sequence 37, App1
90	15	100.0	282	4	US-08-479-453-37	Sequence 37, App1
91	15	100.0	293	4	US-08-701-191A-16	Sequence 16, App1
92	15	100.0	308	4	US-09-333-872A-34	Sequence 34, App1
93	15	100.0	318	4	US-09-199-637A-23	Sequence 23, App1
94	15	100.0	330	2	US-08-712-448-2	Sequence 2, App1
95	15	100.0	333	2	US-08-712-448-1	Sequence 1, App1
96	15	100.0	348	4	US-09-433-702B-263	Sequence 263, App1
97	15	100.0	350	4	US-09-302-769-25	Sequence 25, App1
98	15	100.0	358	1	US-08-239-431A-4	Sequence 4, App1
99	15	100.0	358	2	US-08-463-081B-6	Sequence 6, App1
100	15	100.0	358	2	US-08-461-379A-6	Sequence 6, App1

101	5	33.3	358	2	US-08-462-390B-6	Sequence 6, Appl1	174	5	33.3	1089	1	US-08-168-917-4	Sequence 4, Appl1
102	5	33.3	358	3	US-08-463-074B-6	Sequence 6, Appl1	175	5	33.3	1089	1	US-08-477-329-36	Sequence 36, Appl1
103	5	33.3	358	3	US-08-465-585C-6	Sequence 6, Appl1	176	5	33.3	1089	2	US-08-475-458-36	Sequence 36, Appl1
104	5	33.3	358	3	US-08-652-446-6	Sequence 6, Appl1	177	5	33.3	1089	2	US-08-460-510-4	Sequence 4, Appl1
105	5	33.3	358	4	US-09-267-423-4	Sequence 2, Appl1	178	5	33.3	1089	2	US-08-460-490-4	Sequence 4, Appl1
106	5	33.3	384	4	US-09-552-322-2	Sequence 2, Appl1	179	5	33.3	1089	3	US-08-980-400-36	Sequence 36, Appl1
107	5	33.3	386	4	US-09-590-326-6	Sequence 6, Appl1	180	5	33.3	1089	3	US-08-462-728-2	Sequence 2, Appl1
108	5	33.3	386	6	5171840-5	Sequence 6, Appl1	181	5	33.3	1089	4	US-09-583-459A-36	Sequence 36, Appl1
109	5	33.3	386	6	5480796-5	Sequence 4, Appl1	182	5	33.3	1089	4	US-09-583-210-36	Sequence 36, Appl1
110	5	33.3	388	1	US-08-261-110A-4	Sequence 4, Appl1	183	5	33.3	1089	4	US-09-583-449A-36	Sequence 36, Appl1
111	5	33.3	403	6	5212296-9	Sequence 4, Appl1	184	5	33.3	1089	4	US-09-435-059-36	Sequence 36, Appl1
112	5	33.3	410	4	US-08-725-758A-4	Sequence 14, Appl1	185	5	33.3	1089	4	US-08-461-917-2	Sequence 2, Appl1
113	5	33.3	421	4	US-09-362-473-14	Sequence 14, Appl1	186	5	33.3	1089	5	PCT-US92-00730-4	Sequence 2, Appl1
114	5	33.3	426	4	US-08-725-758A-2	Sequence 2, Appl1	187	5	33.3	1089	5	PCT-US92-00862-4	Sequence 4, Appl1
115	5	33.3	431	4	US-08-845-258-34	Sequence 34, Appl1	188	5	33.3	1135	2	US-08-574-959A-7	Sequence 7, Appl1
116	5	33.3	431	4	US-08-990-571-34	Sequence 34, Appl1	189	5	33.3	1135	4	US-09-357-014-7	Sequence 7, Appl1
117	5	33.3	431	4	US-08-723-142A-34	Sequence 34, Appl1	190	5	33.3	1138	1	US-08-323-474-8	Sequence 8, Appl1
118	5	33.3	431	4	US-09-528-784A-34	Sequence 34, Appl1	191	5	33.3	1138	2	US-08-469-537A-98	Sequence 98, Appl1
119	5	33.3	438	1	US-08-261-110A-2	Sequence 2, Appl1	192	5	33.3	1138	2	US-08-220-240A-5	Sequence 5, Appl1
120	5	33.3	465	1	US-07-872-678A-47	Sequence 47, Appl1	193	5	33.3	1513	5	PCT-US93-03076-2	Sequence 2, Appl1
121	5	33.3	465	2	US-08-588-983-20	Sequence 20, Appl1	194	5	33.3	1833	3	US-08-479-722B-2	Sequence 2, Appl1
122	5	33.3	465	2	US-08-588-983-22	Sequence 22, Appl1	195	5	33.3	1833	5	PCT-US95-02251-18	Sequence 18, Appl1
123	5	33.3	465	2	US-08-588-976-20	Sequence 20, Appl1	196	5	33.3	2318	5	US-09-091-219-24	Sequence 24, Appl1
124	5	33.3	465	2	US-08-588-976-22	Sequence 22, Appl1	197	4	26.7	4	4	US-08-602-999A-97	Sequence 97, Appl1
125	5	33.3	465	4	US-09-347-878-44	Sequence 44, Appl1	198	4	26.7	4	4	US-08-496-841C-162	Sequence 162, App
126	5	33.3	487	1	US-08-249-112-4	Sequence 4, Appl1	199	4	26.7	4	4	US-08-278-865-97	Sequence 97, Appl1
127	5	33.3	487	5	PCT-US95-06556-4	Sequence 4, Appl1	200	4	26.7	4	4	US-09-500-124-97	Sequence 97, Appl1
128	5	33.3	491	4	US-09-426-568A-2	Sequence 2, Appl1							
129	5	33.3	563	4	US-09-431-470-4	Sequence 2, Appl1							
130	5	33.3	589	1	US-08-399-646-2	Sequence 2, Appl1							
131	5	33.3	589	1	US-08-607-321-2	Sequence 2, Appl1							
132	5	33.3	589	2	US-08-961-240-2	Sequence 2, Appl1							
133	5	33.3	589	2	US-08-605-501-2	Sequence 2, Appl1							
134	5	33.3	596	1	US-08-399-646-12	Sequence 12, Appl1							
135	5	33.3	596	1	US-08-607-321-12	Sequence 12, Appl1							
136	5	33.3	596	2	US-08-961-240-12	Sequence 12, Appl1							
137	5	33.3	596	2	US-08-605-501-12	Sequence 12, Appl1							
138	5	33.3	644	5	PCT-US93-00031-19	Sequence 19, Appl1							
139	5	33.3	644	5	PCT-US93-00031-21	Sequence 21, Appl1							
140	5	33.3	647	4	US-09-009-490A-91	Sequence 91, Appl1							
141	5	33.3	647	4	US-08-482-073-5	Sequence 5, Appl1							
142	5	33.3	647	5	PCT-US93-00031-11	Sequence 11, Appl1							
143	5	33.3	662	1	US-08-261-304-7	Sequence 7, Appl1							
144	5	33.3	668	1	US-08-205-018-2	Sequence 2, Appl1							
145	5	33.3	727	1	US-08-424-424B-2	Sequence 2, Appl1							
146	5	33.3	727	5	PCT-US94-05363A-2	Sequence 2, Appl1							
147	5	33.3	731	2	US-08-696-944-20	Sequence 20, Appl1							
148	5	33.3	734	4	US-08-706-216-2	Sequence 2, Appl1							
149	5	33.3	734	4	US-09-641-741-2	Sequence 2, Appl1							
150	5	33.3	735	5	PCT-US93-00031-13	Sequence 13, Appl1							
151	5	33.3	736	5	PCT-US93-00031-15	Sequence 15, Appl1							
152	5	33.3	739	4	US-08-482-073-6	Sequence 6, Appl1							
153	5	33.3	739	5	PCT-US93-00031-9	Sequence 9, Appl1							
154	5	33.3	740	5	PCT-US93-00031-17	Sequence 17, Appl1							
155	5	33.3	755	4	US-09-342-648-2	Sequence 2, Appl1							
156	5	33.3	838	2	US-08-696-944-19	Sequence 19, Appl1							
157	5	33.3	859	1	US-08-395-580-2	Sequence 2, Appl1							
158	5	33.3	859	5	PCT-US95-02792-2	Sequence 2, Appl1							
159	5	33.3	875	4	US-09-150-460B-7	Sequence 7, Appl1							
160	5	33.3	905	2	US-08-574-959A-9	Sequence 9, Appl1							
161	5	33.3	905	4	US-09-357-014-9	Sequence 9, Appl1							
162	5	33.3	919	2	US-08-588-983-9	Sequence 9, Appl1							
163	5	33.3	919	2	US-08-588-983-12	Sequence 12, Appl1							
164	5	33.3	919	2	US-08-588-976-9	Sequence 9, Appl1							
165	5	33.3	919	2	US-08-588-976-12	Sequence 12, Appl1							
166	5	33.3	940	4	US-09-078-347A-1	Sequence 1, Appl1							
167	5	33.3	940	4	US-09-651-656-101	Sequence 101, App							
168	5	33.3	940	4	US-09-650-855-101	Sequence 101, App							
169	5	33.3	955	1	US-08-006-676B-1	Sequence 1, Appl1							
170	5	33.3	955	1	US-08-282-845-2	Sequence 2, Appl1							
171	5	33.3	955	5	US-08-428-414A-3	Sequence 3, Appl1							
172	5	33.3	955	5	PCT-US94-00324-1	Sequence 1, Appl1							
173	5	33.3	1089	1	US-08-180-195-36	Sequence 36, Appl1							

ALIGNMENTS

RESULT 1
US-08-967-101-170
Sequence 170, Application US/08967101
Patent No. 5840540
GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
APPLICANT: ROMMENS, JOHANNA M
APPLICANT: FRASER, PAUL E
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 183
CORRESPONDENCE ADDRESS:
ADDRESSER: TESTA, HURWITZ & THIBEAULT
STREET: High Street Tower - 125 High Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967,101
FILING DATE: 10-NOV-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/592,541
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pletcher, Edmund R.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 170:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: peptide
US-08-967-101-170

Query Match
Best Local Similarity 100.0%; Score 15; DB 2; Length 15;
Pred. No. 1e-09;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLPHRSTPESRAA 15
Db 1 SHLPHRSTPESRAA 15

RESULT 2
US-08-592-541-170
Sequence 170, Application US/08592541
Patent No. 5986054
GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
APPLICANT: ROMMENS, JOHANNA M
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 183
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: High Street Tower - 125 High Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,541
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ. ID NO: 170:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-592-541-170

Query Match
Best Local Similarity 100.0%; Score 15; DB 2; Length 15;
Pred. No. 1e-09;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLPHRSTPESRAA 15
Db 1 SHLPHRSTPESRAA 15

RESULT 3
US-09-124-698-170
Sequence 170, Application US/09124698
Patent No. 6117978
GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
APPLICANT: ROMMENS, JOHANNA M
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 183

CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: High Street Tower - 125 High Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/124,698
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION NUMBER: 08/592,541
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ. ID NO: 170:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-124-698-170

Query Match
Best Local Similarity 100.0%; Score 15; DB 3; Length 15;
Pred. No. 1e-09;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLPHRSTPESRAA 15
Db 1 SHLPHRSTPESRAA 15

RESULT 4
US-09-127-480-170
Sequence 170, Application US/09127480
Patent No. 6194153
GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
APPLICANT: ROMMENS, JOHANNA M
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 183
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: High Street Tower - 125 High Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/127,480
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION NUMBER: US/08/592,541
FILING DATE:
ATTORNEY/AGENT INFORMATION:

NAME: Pitcher, Edmund R.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 170:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-127-480-170

Query Match 100.0%; Score 15; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 1e-09;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPBRSPTESRAA 15
|||||

Db 1 SHLGPBRSPTESRAA 15

RESULT 5
US-08-496-841C-167
Sequence 167, Application US/08496841C
Patent No. 6210919
GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
ROMMENS, JOHANNA M
FRASER, PAUL E

TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TO ALZHEIMER'S DISEASE

NUMBER OF SEQUENCES: 175

CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby & Darby, PC
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10022

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/496,841C
FILING DATE: 28-Jun-1995
CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Paul F. Fehner, Ph.D.
REGISTRATION NUMBER: 35,135
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 527-7700
TELEFAX: (212) 753-6237

INFORMATION FOR SEQ ID NO: 167:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 167:
US-08-496-841C-167

Query Match 100.0%; Score 15; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 1e-09;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPBRSPTESRAA 15
|||||

Db 1 SHLGPBRSPTESRAA 15

RESULT 6
US-09-124-523-170
Sequence 170, Application US/09124523
Patent No. 6395960

GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
ROMMENS, JOHANNA M
FRASER, PAUL E
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 183
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: High Street Tower - 125 High Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/124,523
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/592,541
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 170:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-124-523-170

Query Match 100.0%; Score 15; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 1e-09;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPBRSPTESRAA 15
|||||

Db 1 SHLGPBRSPTESRAA 15

RESULT 7
US-08-875-972-4
Sequence 4, Application US/08875972
Patent No. 5985564
GENERAL INFORMATION:
APPLICANT: Huntington Potter and Jinhue Li
TITLE OF INVENTION: ASSAY FOR IDENTIFYING GENES CAUSING
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
STREET: Two Mallitia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: USA
ZIP: 02173-4799

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/875,972
FILING DATE: 08-AUG-97
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/002,448
FILING DATE: 16-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: Granahan Esq., Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: HU95-03PA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 861-6240
TELEFAX: (781) 861-9540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 407 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-875-972-4

Query Match 100.0%; Score 15; DB 2; Length 407;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGPHRSTPESRAA 15
|||||

Db 286 SHLGPHRSTPESRAA 300

RESULT 8
US-08-788-231A-15
Sequence 15, Application US/08788231A
Patent No. 6019974
GENERAL INFORMATION:
APPLICANT: L'Hernault, Steven W.
TITLE OF INVENTION: SPE-4 RELATED PEPTIDES, ANTIBODIES AND
TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/788,231A
FILING DATE: 24-JAN-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/010,672
FILING DATE: 26-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ferber, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 60-95
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 462 amino acids
TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: not relevant
MOLECULE TYPE: protein
HYPOTHETICAL: YES
US-08-788-231A-15

Query Match 100.0%; Score 15; DB 3; Length 462;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGPHRSTPESRAA 15
|||||

Db 341 SHLGPHRSTPESRAA 355

RESULT 9
US-08-670-479-18
Sequence 18, Application US/08670479
Patent No. 5973133
GENERAL INFORMATION:
APPLICANT: Hardy, John A.
APPLICANT: Goate, Allison M.
TITLE OF INVENTION: MUTANT S182 GENES
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: U.S.A.
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/670,479
FILING DATE: 26-JUN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/001,500
FILING DATE: 18-JUL-1996
APPLICATION NUMBER: 60/001,800
FILING DATE: 02-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: Han, William T.
REGISTRATION NUMBER: 34,344
REFERENCE/DOCKET NUMBER: P50361
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5219
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 463 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
US-08-670-479-18

Query Match 100.0%; Score 15; DB 2; Length 463;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGPHRSTPESRAA 15
|||||

Db 342 SHLGPHRSTPESRAA 356

RESULT 10
US-08-670-964-4
Sequence 4, Application US/08670964
Patent No. 6010874
GENERAL INFORMATION:
APPLICANT: Hardy, John A.
TITLE OF INVENTION: EARLY ONSET ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smithline Beecham Corporation
STREET: 709 Swedeland Road - UW2220; P.O. Box 15
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/670,964
FILING DATE: 26-JUN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/001,142
FILING DATE: 13-JUL-1995
APPLICATION NUMBER: 60/001,501
FILING DATE: 18-JUL-1995
ATTORNEY/AGENT INFORMATION:
NAME: Han, William T
REGISTRATION NUMBER: 34,344
REFERENCE/DOCKET NUMBER: P50358
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5219
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 463 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-670-964-4

Query Match 100.0%; Score 15; DB 3; Length 463;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHGPHRSTPESRAA 15
|||||
DB 342 SHGPHRSTPESRAA 356

RESULT 11
US-08-888-077A-4
Sequence 4, Application US/08888077A
Patent No. 6020143
GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
APPLICANT: ROMMENS, JOHANNA M
APPLICANT: FRASER, PAUL E
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TO ALZHEIMER'S DISEASE AND USES THEREFOR.
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: LERNER, DAVID, LITTEMBERG, KRUMHOLZ & MENTLIK
STREET: 600 SOUTH AVENUE WEST
CITY: WESTFIELD
STATE: NJ

COUNTRY: USA
ZIP: 07090-1497
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/888,077A
FILING DATE: 03-JUL-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/592,541
FILING DATE: 26-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: PALISI, THOMAS M
REGISTRATION NUMBER: 36,629
REFERENCE/DOCKET NUMBER: SCHERING 3.0-017 CIP CIP CIP IV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 654-5000
TELEFAX: (908) 654-7866
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 463 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-888-077A-4

Query Match 100.0%; Score 15; DB 3; Length 463;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHGPHRSTPESRAA 15
|||||
DB 342 SHGPHRSTPESRAA 356

RESULT 12
US-08-788-231A-17
Sequence 17, Application US/08788231A
Patent No. 6019974
GENERAL INFORMATION:
APPLICANT: L'Hernault, Steven W.
TITLE OF INVENTION: SPE-4 RELATED PEPTIDES, ANTIBODIES AND
METHODS
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/788,231A
FILING DATE: 24-JAN-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/010,672
FILING DATE: 26-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Feider, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 60-95
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089

INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 465 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: protein
HYPOTHETICAL: YES
US-08-788-231A-17

Query Match 100.0%; Score 15; DB 3; Length 465;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPBRSPTESRAA 15
|||||
DB 346 SHLGPBRSPTESRAA 358

RESULT 13

US-08-967-101-2
Sequence 2, Application US/08967101
Patent No. 5840540

GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
APPLICANT: ROMMENS, JOHANNA M
APPLICANT: FRASER, PAUL E
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 183
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: High Street Tower - 125 High Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967,101
FILING DATE: 10-NOV-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/592,541

1-26-94

ATTORNEY/AGENT INFORMATION:
NAME: Pletcher, Edmund R.
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-967-101-2

Query Match 100.0%; Score 15; DB 2; Length 467;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPBRSPTESRAA 15
|||||
DB 346 SHLGPBRSPTESRAA 360

RESULT 14

US-08-967-101-4
Sequence 4, Application US/08967101
Patent No. 5840540

GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
APPLICANT: ROMMENS, JOHANNA M
APPLICANT: FRASER, PAUL E
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 183
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: High Street Tower - 125 High Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967,101
FILING DATE: 10-NOV-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/592,541

1-26-94

ATTORNEY/AGENT INFORMATION:
NAME: Pletcher, Edmund R.
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-967-101-4

Query Match 100.0%; Score 15; DB 2; Length 467;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPBRSPTESRAA 15
|||||
DB 346 SHLGPBRSPTESRAA 360

RESULT 15

US-08-967-101-134
Sequence 134, Application US/08967101
Patent No. 5840540

GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
APPLICANT: ROMMENS, JOHANNA M
APPLICANT: FRASER, PAUL E
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 183
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: High Street Tower - 125 High Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967,101
FILING DATE: 10-NOV-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/592,541
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 134:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-967-101-134

Query Match 100.0%; Score 15; DB 2; Length 467;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPFRSTPESRAA 15
DB 346 SHLGPFRSTPESRAA 360
|||||

RESULT 16
US-08-592-541-2
Sequence 2, Application US/08592541
Patent No. 5986054
GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
APPLICANT: ROMMENS, JOHANNA M
APPLICANT: FRASER, PAUL E
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 183
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: High Street Tower - 125 High Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,541
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-592-541-2

Query Match 100.0%; Score 15; DB 2; Length 467;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPFRSTPESRAA 15
DB 346 SHLGPFRSTPESRAA 360
|||||

RESULT 17
US-08-592-541-4
Sequence 4, Application US/08592541
Patent No. 5986054
GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
APPLICANT: ROMMENS, JOHANNA M
APPLICANT: FRASER, PAUL E
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 183
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: High Street Tower - 125 High Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,541
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-592-541-4

Query Match 100.0%; Score 15; DB 2; Length 467;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPFRSTPESRAA 15
DB 346 SHLGPFRSTPESRAA 360
|||||

RESULT 18
US-08-592-541-134
Sequence 134, Application US/08592541
Patent No. 5986054
GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
APPLICANT: ROMMENS, JOHANNA M
APPLICANT: FRASER, PAUL E
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 183
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: High Street Tower - 125 High Street

CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,541
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 134:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-592-541-134

Query Match 100.0%; Score 15; DB 2; Length 467;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGPHRSTPESRAA 15
Db 346 SHLGPHRSTPESRAA 360

RESULT 19
US-08-923-454A-10
Sequence 10, Application US/08923454A
Patent No. 6004794
GENERAL INFORMATION:
APPLICANT: Creasy, Caretha
APPLICANT: Livl, George
APPLICANT: Karran, Eric
APPLICANT: Clinkenbeard, Helen
APPLICANT: Browne, Michael
APPLICANT: Southan, Christopher
TITLE OF INVENTION: HUMAN SERINE PROTEASE
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/923,454A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/025436
FILING DATE: 06-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Baumelster, Kirk
REGISTRATION NUMBER: 33,833
REFERENCE/DOCKET NUMBER: P50547
TELECOMMUNICATION INFORMATION:

TELEPHONE: 610-270-5096
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
US-08-923-454A-10

Query Match 100.0%; Score 15; DB 3; Length 467;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGPHRSTPESRAA 15
Db 346 SHLGPHRSTPESRAA 360

RESULT 20
US-08-670-964-2
Sequence 2, Application US/08670964
Patent No. 6010874
GENERAL INFORMATION:
APPLICANT: Hardy, John A.
TITLE OF INVENTION: EARLY ONSET ALZHEIMER'S DISEASE
TITLE OF INVENTION: GENE AND GENE PRODUCTS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road - UW2220; P.O. Box 15
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/670,964
FILING DATE: 26-JUN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/001,142
FILING DATE: 18-JUL-1995
APPLICATION NUMBER: 60/001,501
FILING DATE: 18-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Han, William T
REGISTRATION NUMBER: 34,344
REFERENCE/DOCKET NUMBER: P50358
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5219
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-670-964-2
Query Match 100.0%; Score 15; DB 3; Length 467;

Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGPBRSPTESRAA 15
Db 346 SHLGPBRSPTESRAA 360

RESULT 21

US-08-888-077A-2
Sequence 2, Application US/08888077A
Patent No. 6020143
GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
APPLICANT: ROMMENS, JOHANNA M
APPLICANT: FRASER, PAUL E
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TITLE OF INVENTION: TO ALZHEIMER'S DISEASE AND USES THEREFOR.
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: LERNER, DAVID, LITTENBERG, KRUMHOLZ & MENTLIK
STREET: 600 SOUTH AVENUE WEST
CITY: WESTFIELD
STATE: NJ
COUNTRY: USA
ZIP: 07090-1497

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/888,077A
FILING DATE: 03-JUL-1997

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/592,541
FILING DATE: 26-JAN-1996

ATTORNEY/AGENT INFORMATION:

NAME: PALISI, THOMAS M

REGISTRATION NUMBER: 36,629

REFERENCE/DOCKET NUMBER: SCHERING 3.0-017 CIP CIP IV

TELECOMMUNICATION INFORMATION:

TELEPHONE: (908) 654-5000

TELEFAX: (908) 654-7866

INFORMATION FOR SEQ. ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 467 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-888-077A-2

Query Match

Best Local Similarity 100.0%; Score 15; DB 3; Length 467;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGPBRSPTESRAA 15
Db 346 SHLGPBRSPTESRAA 360

RESULT 22

US-08-888-077A-17

Sequence 17, Application US/08888077A

Patent No. 6020143

GENERAL INFORMATION:

APPLICANT: ST. GEORGE-HYSLOP, PETER H

APPLICANT: ROMMENS, JOHANNA M

APPLICANT: FRASER, PAUL E

TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED

TITLE OF INVENTION: TO ALZHEIMER'S DISEASE AND USES THEREFOR.

NUMBER OF SEQUENCES: 41

CORRESPONDENCE ADDRESS:

ADDRESSEE: LERNER, DAVID, LITTENBERG, KRUMHOLZ & MENTLIK
STREET: 600 SOUTH AVENUE WEST
CITY: WESTFIELD
STATE: NJ
COUNTRY: USA
ZIP: 07090-1497

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/888,077A

FILING DATE: 03-JUL-1997

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/592,541

FILING DATE: 26-JAN-1996

ATTORNEY/AGENT INFORMATION:

NAME: PALISI, THOMAS M

REGISTRATION NUMBER: 36,629

REFERENCE/DOCKET NUMBER: SCHERING 3.0-017 CIP CIP IV

TELECOMMUNICATION INFORMATION:

TELEPHONE: (908) 654-5000

TELEFAX: (908) 654-7866

INFORMATION FOR SEQ. ID NO: 17:

SEQUENCE CHARACTERISTICS:

LENGTH: 467 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-888-077A-17

Query Match

Best Local Similarity 100.0%; Score 15; DB 3; Length 467;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGPBRSPTESRAA 15
Db 346 SHLGPBRSPTESRAA 360

RESULT 23

US-09-124-698-2

Sequence 2, Application US/09124698

Patent No. 6117978

GENERAL INFORMATION:

APPLICANT: ST. GEORGE-HYSLOP, PETER H

APPLICANT: ROMMENS, JOHANNA M

APPLICANT: FRASER, PAUL E

TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED

TITLE OF INVENTION: TO ALZHEIMER'S DISEASE

NUMBER OF SEQUENCES: 183

CORRESPONDENCE ADDRESS:

ADDRESSEE: TESTA, HURWITZ & THIBEAULT

STREET: High Street Tower - 125 High Street

CITY: Boston

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02110

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/124,698

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/592,541

FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Pltcher, Edmund R.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-124-698-2

Query Match 100.0%; Score 15; DB 3; Length 467;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGPHRSTPESRAA 15
Db 346 SHLGPHRSTPESRAA 360

RESULT 24
US-09-124-698-4
Sequence 4, Application US/09124698
Patent No. 6117978

GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
APPLICANT: ROMMENS, JOHANNA M
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 183
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: High Street Tower - 125 High Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/124,698
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/592,541
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pltcher, Edmund R.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-124-698-4

Query Match 100.0%; Score 15; DB 3; Length 467;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGPHRSTPESRAA 15
Db 346 SHLGPHRSTPESRAA 360

Db 346 SHLGPHRSTPESRAA 360

RESULT 25
US-09-124-698-134
Sequence 134, Application US/09124698
Patent No. 6117978

GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
APPLICANT: ROMMENS, JOHANNA M
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 183
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: High Street Tower - 125 High Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/124,698
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/592,541
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pltcher, Edmund R.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 134:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-124-698-134

Query Match 100.0%; Score 15; DB 3; Length 467;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGPHRSTPESRAA 15
Db 346 SHLGPHRSTPESRAA 360

RESULT 26
US-09-127-480-2
Sequence 2, Application US/09127480
Patent No. 6194153
GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
APPLICANT: ROMMENS, JOHANNA M
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 183
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: High Street Tower - 125 High Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.

ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/127,480
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/592,541
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-127-480-2

Query Match 100.0%; Score 15; DB 4; Length 467;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPRTSPESRAA 15
|||||

Db 346 SHLGPRTSPESRAA 360

RESULT 27
US-09-127-480-4
Sequence 4, Application US/09127480
Patent No. 6194153
GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
APPLICANT: ROMMENS, JOHANNA M
APPLICANT: FRASER, PAUL E
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 183
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: High Street Tower - 125 High Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/127,480
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/592,541
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:

LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-127-480-4

Query Match 100.0%; Score 15; DB 4; Length 467;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPRTSPESRAA 15
|||||

Db 346 SHLGPRTSPESRAA 360

RESULT 28
US-09-127-480-134
Sequence 134, Application US/09127480
Patent No. 6194153
GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
APPLICANT: ROMMENS, JOHANNA M
APPLICANT: FRASER, PAUL E
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 183
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: High Street Tower - 125 High Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/127,480
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/592,541
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 134:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-127-480-134

Query Match 100.0%; Score 15; DB 4; Length 467;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPRTSPESRAA 15
|||||

Db 346 SHLGPRTSPESRAA 360

RESULT 29
US-08-496-841C-2
Sequence 2, Application US/08496841C
Patent No. 6210919

GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
ROMMENS, JOHANNA M
FRASER, PAUL E
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 175
CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby & Darby, PC
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/496,841C
FILING DATE: 28-Jun-1995
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Paul F. Fehner, Ph.D.
REGISTRATION NUMBER: 35,135
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 527-7700
TELEFAX: (212) 753-6237
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-08-496-841C-2

Query Match 100.0%; Score 15; DB 4; Length 467;
Best Local Similarity 100.0%; Pred. No. 2e-08; 0; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 0;

QY 1 SHLPHRSTPESRAA 15
|||||

Db 346 SHLPHRSTPESRAA 360

RESULT 30
US-08-496-841C-4
Sequence 4, Application US/08496841C
Patent No. 6210919
GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
ROMMENS, JOHANNA M
FRASER, PAUL E
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 175
CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby & Darby, PC
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/496,841C

FILING DATE: 28-Jun-1995
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Paul F. Fehner, Ph.D.
REGISTRATION NUMBER: 35,135
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 527-7700
TELEFAX: (212) 753-6237
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-08-496-841C-4

Query Match 100.0%; Score 15; DB 4; Length 467;
Best Local Similarity 100.0%; Pred. No. 2e-08; 0; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 0;

QY 1 SHLPHRSTPESRAA 15
|||||

Db 346 SHLPHRSTPESRAA 360

RESULT 31
US-08-496-841C-134
Sequence 134, Application US/08496841C
Patent No. 6210919
GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
ROMMENS, JOHANNA M
FRASER, PAUL E
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 175
CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby & Darby, PC
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/496,841C
FILING DATE: 28-Jun-1995
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Paul F. Fehner, Ph.D.
REGISTRATION NUMBER: 35,135
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 527-7700
TELEFAX: (212) 753-6237
INFORMATION FOR SEQ ID NO: 134:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 134:
US-08-496-841C-134

Query Match 100.0%; Score 15; DB 4; Length 467;
Best Local Similarity 100.0%; Pred. No. 2e-08; 0; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 0;

QY 1 SHLGPBRSPTESRAA 15
DB 346 SHLGPBRSPTESRAA 360

RESULT 32

US-08-496-841C-136
Sequence 136, Application US/08496841C
Patent No. 6210919
GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
ROMMENS, JOHANNNA M
FRASER, PAUL E
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 175
CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby & Darby, PC
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/496,841C
FILING DATE: 28-Jun-1995
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Paul F. Fehlner, Ph.D.
REGISTRATION NUMBER: 35,135
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 527-7700
TELEFAX: (212) 753-6237
INFORMATION FOR SEQ ID NO: 136:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 136:
US-08-496-841C-136
Query Match 100.0%; Score 15; DB 4; Length 467;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPBRSPTESRAA 15
DB 346 SHLGPBRSPTESRAA 360

RESULT 33

US-08-706-344C-2
Sequence 2, Application US/08706344C
Patent No. 6248555
GENERAL INFORMATION:
APPLICANT: TANZI, RUDOLPH
APPLICANT: WASCO, WILMA
TITLE OF INVENTION: Genetic Alterations Related To Familial
TITLE OF INVENTION: Alzheimer's Disease
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVENUE, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA

ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/706,344C
FILING DATE: 30-AUG-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/003,054
FILING DATE: 31-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: KIM, JUDITH U.
REGISTRATION NUMBER: 40,679
REFERENCE/DOCKET NUMBER: 0609.4180001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2540
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-706-344C-2

QY 1 SHLGPBRSPTESRAA 15
DB 346 SHLGPBRSPTESRAA 360

RESULT 34

US-08-706-344C-4
Sequence 4, Application US/08706344C
Patent No. 6248555
GENERAL INFORMATION:
APPLICANT: TANZI, RUDOLPH
APPLICANT: WASCO, WILMA
TITLE OF INVENTION: Genetic Alterations Related To Familial
TITLE OF INVENTION: Alzheimer's Disease
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVENUE, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/706,344C
FILING DATE: 30-AUG-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/003,054
FILING DATE: 31-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: KIM, JUDITH U.
REGISTRATION NUMBER: 40,679
REFERENCE/DOCKET NUMBER: 0609.4180001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540

;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 467 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-706-344C-4

Query Match 100.0%; Score 15; DB 4; Length 467;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGPSTPESRAA 15
Db 346 SHLGPSTPESRAA 360

RESULT 35
US-08-706-344C-28
; Sequence 28, Application US/08706344C
; Patent No. 6248555
; GENERAL INFORMATION:
; APPLICANT: TANZI, RUDOLPH
; APPLICANT: MASCO, WILMA
; TITLE OF INVENTION: Genetic Alterations Related To Familial
; TITLE OF INVENTION: Alzheimer's Disease
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/706,344C
; FILING DATE: 30-AUG-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/003,054
; FILING DATE: 31-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: KIM, JUDITH U.
; REGISTRATION NUMBER: 40,679
; REFERENCE/DOCKET NUMBER: 0609,4180001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-706-344C-28

Query Match 100.0%; Score 15; DB 4; Length 467;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGPSTPESRAA 15
Db 346 SHLGPSTPESRAA 360

RESULT 36
US-08-706-344C-30
; Sequence 30, Application US/08706344C

;; Patent No. 6248555
;; GENERAL INFORMATION:
;; APPLICANT: TANZI, RUDOLPH
;; APPLICANT: MASCO, WILMA
;; TITLE OF INVENTION: Genetic Alterations Related To Familial
;; TITLE OF INVENTION: Alzheimer's Disease
;; NUMBER OF SEQUENCES: 32
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
;; STREET: 1100 NEW YORK AVENUE, SUITE 600
;; CITY: WASHINGTON
;; STATE: DC
;; COUNTRY: USA
;; ZIP: 20005-3934
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: 60/003,054
;; FILING DATE: 31-AUG-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: KIM, JUDITH U.
;; REGISTRATION NUMBER: 40,679
;; REFERENCE/DOCKET NUMBER: 0609,4180001
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-371-2600
;; TELEFAX: 202-371-2540
;; INFORMATION FOR SEQ ID NO: 30:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 467 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-706-344C-30

Query Match 100.0%; Score 15; DB 4; Length 467;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGPSTPESRAA 15
Db 346 SHLGPSTPESRAA 360

RESULT 37
US-08-706-344C-32
; Sequence 32, Application US/08706344C
; Patent No. 6248555
; GENERAL INFORMATION:
; APPLICANT: TANZI, RUDOLPH
; APPLICANT: MASCO, WILMA
; TITLE OF INVENTION: Genetic Alterations Related To Familial
; TITLE OF INVENTION: Alzheimer's Disease
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/706,344C
FILING DATE: 30-AUG-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/003,054
FILING DATE: 31-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: KIM, JUDITH U.
REGISTRATION NUMBER: 40,679
REFERENCE/DOCKET NUMBER: 0609.4180001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-706-344C-32

Query Match 100.0%; Score 15; DB 4; Length 467;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPHRSTPESRAA 15
|||||
DB 346 SHLGPHRSTPESRAA 360

RESULT 38
US-08-832-867-3
Sequence 3, Application US/08832867C
Patent No. 6376239
GENERAL INFORMATION:
APPLICANT: BAUMEISTER, Ralf
TITLE OF INVENTION: DNA MOLECULES COMPRISING A PROMOTER CAPABLE OF
TITLE OF INVENTION: CONFERRING EXPRESSION OF A HETEROLOGOUS DNA SEQUENCE IN
TITLE OF INVENTION: ALL NEURAL CELLS AT ALL STAGES OF DEVELOPMENT IN C.
TITLE OF INVENTION: ELEGANS AND USES THEREOF
FILE REFERENCE: 674503/2004
CURRENT APPLICATION NUMBER: US/08/832,867C
CURRENT FILING DATE: 1997-04-04
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 467
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-08-832-867-3

Query Match 100.0%; Score 15; DB 4; Length 467;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPHRSTPESRAA 15
|||||
DB 346 SHLGPHRSTPESRAA 360

RESULT 39
US-09-227-725A-1
Sequence 1, Application US/09227725A
Patent No. 6383758
GENERAL INFORMATION:
APPLICANT: St. George-Hyslop, Peter H.
APPLICANT: Rommens, Johanna
APPLICANT: Fraser, Paul E.
TITLE OF INVENTION: Alzheimer's Related Proteins and Methods
TITLE OF INVENTION: of Use
FILE REFERENCE: 1034/1F810-US1
CURRENT APPLICATION NUMBER: US/09/227,725A
CURRENT FILING DATE: 1999-01-08

NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 467
TYPE: PRT
ORGANISM: Homo Saplen
US-09-227-725A-1

Query Match 100.0%; Score 15; DB 4; Length 467;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPHRSTPESRAA 15
|||||
DB 346 SHLGPHRSTPESRAA 360

RESULT 40
US-09-124-523-2
Sequence 2, Application US/09124523
Patent No. 6395960
GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
APPLICANT: ROMMENS, JOHANNA M
APPLICANT: FRASER, PAUL E
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: High Street Tower - 125 High Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/124,523
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/592,541
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-124-523-2

Query Match 100.0%; Score 15; DB 4; Length 467;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPHRSTPESRAA 15
|||||
DB 346 SHLGPHRSTPESRAA 360

RESULT 41
US-09-124-523-4
Sequence 4, Application US/09124523

Patent No. 6395960
GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
APPLICANT: ROMMENS, JOHANNA M
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 183
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: High Street tower - 125 High Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/124,523
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/592,541
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-124-523-4
Query Match 100.0%; Score 15; DB 4; Length 467;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SHLGPSTPESRA 15
DB 346 SHLGPSTPESRA 360
RESULT 42
US-09-124-523-134
Sequence 134, Application US/09124523
Patent No. 6395960
GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
APPLICANT: ROMMENS, JOHANNA M
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 183
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: High Street tower - 125 High Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/124,523
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/592,541
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 134:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-124-523-134
Query Match 100.0%; Score 15; DB 4; Length 467;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SHLGPSTPESRA 15
DB 346 SHLGPSTPESRA 360
RESULT 43
US-09-375-318-3
Sequence 3, Application US/09375318
Patent No. 6468791
GENERAL INFORMATION:
APPLICANT: Tanzi, Rudolph E.
Masco, Wilma
Levy-Lahad, Ephrat
Blid, Thomas D.
Galas, David J.
TITLE OF INVENTION: CHROMOSOME 1 GENE AND GENE PRODUCTS RELATED
TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 88
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 701 Fifth Ave, Suite 6300
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/375,318
FILING DATE: 16-Aug-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Verna, James M.
REGISTRATION NUMBER: 33,287
REFERENCE/DOCKET NUMBER: 920010.571C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-375-318-3
Query Match 100.0%; Score 15; DB 4; Length 467;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHLGHRSPTESRAA 15
DB 346 SHLGHRSPTESRAA 360
RESULT 44
US-09-325-932A-57
Sequence 57, Application US/09325932A
Patent No. 6451604
GENERAL INFORMATION:
APPLICANT: Flinn, Barry
TITLE OF INVENTION: Compositions affecting programmed cell
TITLE OF INVENTION: death and their use in the modification of forestry plant develo
FILE REFERENCE: 1022
CURRENT APPLICATION NUMBER: US/09/325,932A
CURRENT FILING DATE: 1999-08-04
NUMBER OF SEQ ID NOS: 206
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 57
LENGTH: 256
TYPE: PRT
ORGANISM: Pinus radiata
US-09-325-932A-57
Query Match 40.0%; Score 6; DB 4; Length 256;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 10 PESRAA 15
DB 184 PESRAA 189
RESULT 45
US-07-857-224B-76
Sequence 76, Application US/07857224B
Patent No. 5958784
GENERAL INFORMATION:
APPLICANT: Benner, Steven A.
TITLE OF INVENTION: Predicting folded structures of proteins
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Steven A. Benner
STREET: Hadlaubstrasse 151
CITY: Zurich
STATE: none
COUNTRY: Switzerland
ZIP: (note: this is an international post code) CH-8092
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette, 1.4 MB storage
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/857,224B
FILING DATE: 03/25/92
CLASSIFICATION: 436
PRIOR APPLICATION DATA: none
TELECOMMUNICATION INFORMATION:
TELEPHONE: (international) 41 1 632 2830
TELEFAX: (international) 41 1 262 2437
TELEX: none
INFORMATION FOR SEQ ID NO: 76:
SEQUENCE CHARACTERISTICS:
LENGTH: 270

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: protein
ORIGINAL SOURCE:
ORGANISM: mouse
FEATURE: Protein kinase; Table 8 Column 87
PUBLICATION INFORMATION:
AUTHORS:
AUTHORS: Hanks, S. K.
AUTHORS: Quinn, A. M.
AUTHORS: Hunter, T.
TITLE: The protein kinase family
JOURNAL: Science
VOLUME: 241
PAGES: 42-52
DATE: 1988
US-07-857-224B-76
Query Match 40.0%; Score 6; DB 2; Length 270;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHLGH 6
DB 59 SHLGH 64
RESULT 46
US-08-844-055-2
Sequence 2, Application US/08844055
Patent No. 5747313
GENERAL INFORMATION:
APPLICANT: Lawlor, Elizabeth
TITLE OF INVENTION: No. 5747313e1 Compounds
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/844,055
FILING DATE: 18-APR-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9607993.4
FILING DATE: 18-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimm1, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31457-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 369 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-844-055-2
Query Match 40.0%; Score 6; DB 1; Length 369;

Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRAA 15
|||||

Db 171 PESRAA 176

RESULT 47

US-09-006-849-2

; Sequence 2, Application US/09006849

; Patent No. 6071731

; GENERAL INFORMATION:

; APPLICANT: Lawlor, Elizabeth

; TITLE OF INVENTION: No. 6071731el Compounds

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SmithKline Beecham Corporation

; STREET: 709 Swedeland Road

; CITY: King of Prussia

; STATE: PA

; COUNTRY: USA

; ZIP: 19406-0939

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; OPERATING SYSTEM: DOS

; SOFTWARE: FASTSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/006,849

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/844,055

; FILING DATE: 18-APR-1997

; APPLICATION NUMBER: 9607993.4

; FILING DATE: 18-APR-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Gimmil, Edward R

; REGISTRATION NUMBER: 38,891

; REFERENCE/DOCKET NUMBER: P31457-1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 610-270-4478

; TELEFAX: 610-270-5090

; TELEX:

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 369 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-09-006-849-2

Query Match 40.0%; Score 6; DB 3;

Best Local Similarity 100.0%; Pred. No. 34;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRAA 15
|||||

Db 171 PESRAA 176

RESULT 48

US-08-906-744A-2

; Sequence 2, Application US/08906744A

; Patent No. 5795758

; GENERAL INFORMATION:

; APPLICANT: Gentry, Daniel

; APPLICANT: Greenwood, Rebecca

; APPLICANT: Lawlor, Elizabeth

; TITLE OF INVENTION: NOVEL HISS

; NUMBER OF SEQUENCES: 2

US-08-906-744A-2

Query Match 40.0%; Score 6; DB 3;

Best Local Similarity 100.0%; Pred. No. 34;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRAA 15
|||||

Db 171 PESRAA 176

CORRESPONDENCE ADDRESS:

ADDRESSEE: SmithKline Beecham Corporation

STREET: 709 Swedeland Road

CITY: King of Prussia

STATE: PA

COUNTRY: USA

ZIP: 19406-0939

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/906,744A

FILING DATE: 06-AUG-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/844,055

FILING DATE: 18-APR-1997

APPLICATION NUMBER: 9607993.4

FILING DATE: 18-APR-1996

ATTORNEY/AGENT INFORMATION:

NAME: Gimmil, Edward R

REGISTRATION NUMBER: 38,891

REFERENCE/DOCKET NUMBER: P31457-1/1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 610-270-4478

TELEFAX: 610-270-5090

TELEX:

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 429 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-906-744A-2

Query Match 40.0%; Score 6; DB 1;

Best Local Similarity 100.0%; Pred. No. 38;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRAA 15
|||||

Db 171 PESRAA 176

RESULT 49

US-09-093-134-2

; Sequence 2, Application US/09093134

; Patent No. 6040162

; GENERAL INFORMATION:

; APPLICANT: Gentry, Daniel

; APPLICANT: Greenwood, Rebecca

; APPLICANT: Lawlor, Elizabeth

; TITLE OF INVENTION: NOVEL HISS

; NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: SmithKline Beecham Corporation

STREET: 709 Swedeland Road

CITY: King of Prussia

STATE: PA

COUNTRY: USA

ZIP: 19406-0939

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/093,134

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/906,744
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/844,055
FILING DATE: 18-APR-1997
APPLICATION NUMBER: 9607993.4
FILING DATE: 18-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ghiml, Edward R.
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31457-1/1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 429 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-093-134-2

Query Match 40.0%; Score 6; DB 3; Length 429;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRAA 15
|||||
DB 171 PESRAA 176

RESULT 50
US-08-810-116-11
Sequence 11, Application US/08810116
Patent No. 5766860
GENERAL INFORMATION:
APPLICANT: Terman, Bruce I.
APPLICANT: Carrion, Miguel E.
TITLE OF INVENTION: Identification of a No. 5766860el Human Growth
TITLE OF INVENTION: Factor Receptor
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: American Cyanamid Company
STREET: One Cyanamid Plaza
CITY: Wayne
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07470
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/810,116
FILING DATE: 25-FEB-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/930,548
FILING DATE: 23-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Gordon, Alan M.
REGISTRATION NUMBER: 30,637
REFERENCE/DOCKET NUMBER: 31,298-01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-831-3244
TELEFAX: 201-831-3305
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 566 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
AUTHORS: Gronwald, R., et al.
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
VOLUME: 85
PAGES: 3435-3439
DATE: 1988
US-08-810-116-11

Query Match 40.0%; Score 6; DB 1; Length 566;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGPH 6
|||||
DB 103 SHLGPH 108

RESULT 51
US-07-930-548A-11
Sequence 11, Application US/07930548A
Patent No. 5861301
GENERAL INFORMATION:
APPLICANT: Terman, Bruce I.
APPLICANT: Carrion, Miguel E.
TITLE OF INVENTION: Identification of a No. 5861301el Human Growth
TITLE OF INVENTION: Factor Receptor
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: American Cyanamid Company
STREET: One Cyanamid Plaza
CITY: Wayne
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07470
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/930,548A
FILING DATE: 23-NOV-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gordon, Alan M.
REGISTRATION NUMBER: 30,637
REFERENCE/DOCKET NUMBER: 31,298-01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-831-3244
TELEFAX: 201-831-3305
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 566 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
AUTHORS: Gronwald, R., et al.
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
VOLUME: 85
PAGES: 3435-3439
DATE: 1988
US-07-930-548A-11

Query Match 40.0%; Score 6; DB 2; Length 566;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGPH 6
|||||

Db 103 SHLGP 108

RESULT 52

US-07-803-633A-13

; Sequence 13, Application US/07803633A
; Patent No. 5369025

GENERAL INFORMATION:

APPLICANT: NAZERIAN, Keyvan

APPLICANT: LEE, Lucy F.

APPLICANT: VANAGIDA, No. 5369025oru

APPLICANT: OGAWA, Ryohei

APPLICANT: LI, Yi

TITLE OF INVENTION: RECOMBINANT FOWLPOX VACCINE FOR
PROTECTION AGAINST MARX'S DISEASE

NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:

ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH

STREET: 301 No. 5369025th Washington Street

CITY: Falls Church

STATE: Virginia

COUNTRY: USA

ZIP: 22040-0747

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentln Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/803,633A

FILING DATE: 19911210

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Murphy Jr., Gerald M.

REGISTRATION NUMBER: 28,977

REFERENCE/DOCKET NUMBER: 1644-103P

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 241-1300

TELEFAX: (703) 241-2848

TELEX: 248345

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 865 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein

US-07-803-633A-13

QY 9 TPESRA 14

Db 174 TPESRA 179

RESULT 53

US-08-180-195-2

; Sequence 2, Application US/08180195
; Patent No. 5567584

GENERAL INFORMATION:

APPLICANT: Siedziewski Ph.D., Andrzej Z

APPLICANT: Bell, Lillian A.

APPLICANT: Kindsvogel Ph.D., Wayne R.

TITLE OF INVENTION: METHODS OF PRODUCING SECRETED RECEPTOR ANALOGS

TITLE OF INVENTION: AND BIOLOGICALLY ACTIVE DIMERIZED POLYPEPTIDE

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed and Berry

STREET: 6300 Columbia Center

CITY: Seattle

STATE: WA
COUNTRY: USA
ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentln Release #1.24

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/180,195

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/634,510

FILING DATE:

APPLICATION NUMBER: US 07/146,877

FILING DATE: 22-JAN-1988

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/347,291

FILING DATE: 02-MAY-1989

ATTORNEY/AGENT INFORMATION:

NAME: Maki J.D., David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 990008,446C3

TELECOMMUNICATION INFORMATION:

TELEPHONE: 206-622-4900

TELEFAX: 206-682-6031

TELEX: 3723836

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1106 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-180-195-2

QY 1 SHLGP 6

Db 656 SHLGP 661

RESULT 54

US-08-168-917-2

; Sequence 2, Application US/08168917
; Patent No. 5686572

GENERAL INFORMATION:

APPLICANT: Wolf, David

APPLICANT: Tomlinson, James E.

APPLICANT: Fretto, Larry J.

APPLICANT: Giese, Neill A.

APPLICANT: Escobedo, Jaime A.

APPLICANT: Williams, Lewis T.

TITLE OF INVENTION: DOMAINS OF EXTRACELLULAR REGION OF HUMAN

TITLE OF INVENTION: PLATELET-DERIVED GROWTH FACTOR RECEPTOR POLYPEPTIDES

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend

STREET: Steuart Street Tower, 20th Floor \ One Market

CITY: San Francisco

STATE: California

COUNTRY: US

ZIP: 94105

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentln Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/168,917
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/650,793
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: 12418-14
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1106 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-168-917-2

Query Match 40.0%; Score 6; DB 1; Length 1106;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGP 6
DB 656 SHLGP 661

RESULT 55

US-08-477-329-2
Sequence 2, Application US/08477329
Patent No. 5750375

GENERAL INFORMATION:
APPLICANT: Sledziewski Ph.D., Andrzej Z

APPLICANT: Bell, Lillian A.

TITLE OF INVENTION: METHODS OF PRODUCING SECRETED RECEPTOR ANALOGS AND BIOLOGICAL

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:
ADDRESS: Seed and Berry

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: WA

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.24

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,329

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 990008.446C6

TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900

TELEFAX: 206-682-6031

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1106 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-477-329-2

Query Match 40.0%; Score 6; DB 1; Length 1106;

Best Local Similarity 100.0%; Pred. No. 87;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGP 6
DB 656 SHLGP 661

RESULT 56

US-08-475-458-2
Sequence 2, Application US/08475458
Patent No. 5843725

GENERAL INFORMATION:
APPLICANT: Sledziewski Ph.D., Andrzej Z

APPLICANT: Bell, Lillian A.

TITLE OF INVENTION: METHODS OF PRODUCING SECRETED RECEPTOR ANALOGS AND BIOLOGIC

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:
ADDRESS: Seed and Berry

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: WA

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.24

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,458

FILING DATE: 06-JUN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 990008.446D5

TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900

TELEFAX: 206-682-6031

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1106 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-475-458-2

Query Match 40.0%; Score 6; DB 2; Length 1106;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGP 6
DB 656 SHLGP 661

RESULT 57

US-08-460-510-2
Sequence 2, Application US/08460510
Patent No. 5872218

GENERAL INFORMATION:
APPLICANT: Wolf, David

APPLICANT: Tomlinson, James E.

APPLICANT: Pietro, Larry J.

APPLICANT: Giese, Neill A.

APPLICANT: Escobedo, Jaime A.

APPLICANT: Williams, Lewis T.

TITLE OF INVENTION: DOMAINS OF EXTRACELLULAR REGION OF HUMAN

PLATELET-DERIVED GROWTH FACTOR RECEPTOR POLYPEPTIDES

NUMBER OF SEQUENCES: 23

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: TOWNSEND and TOWNSEND and CREW
;; STREET: One Market Plaza, Stewart Street Tower, Suite 2000
;; CITY: San Francisco
;; STATE: California
;; COUNTRY: US
;; ZIP: 94105
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/460,510
;; FILING DATE: 02-JUN-1995
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Dow, Karen B.
;; REGISTRATION NUMBER: 29,684
;; REFERENCE/DOCKET NUMBER: 012418-001430
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 326-2400
;; TELEFAX: (415) 326-2422
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1106 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-460-510-2

Query Match
Best Local Similarity 40.0%; Score 6; DB 2; Length 1106;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGP 6
Db 656 SHLGP 661

RESULT 58
US-08-460-490-2
; Sequence 2, Application US/08460490
; Patent No. 5891652
; GENERAL INFORMATION:
; APPLICANT: Wolf, David
; APPLICANT: Tomlinson, James E.
; APPLICANT: Fretto, Larry J.
; APPLICANT: Gleese, Neil A.
; APPLICANT: Escobedo, Jaime A.
; APPLICANT: Williams, Lewis T.
; TITLE OF INVENTION: DOMAINS OF EXTRACELLULAR REGION OF HUMAN
; TITLE OF INVENTION: PLATELET-DERIVED GROWTH FACTOR RECEPTOR POLYPEPTIDES
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW
; STREET: One Market Plaza, Stewart Street Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,490
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684

;; REFERENCE/DOCKET NUMBER: 012418-001420
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 326-2400
;; TELEFAX: (415) 326-2422
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1106 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-460-490-2

Query Match
Best Local Similarity 40.0%; Score 6; DB 2; Length 1106;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGP 6
Db 656 SHLGP 661

RESULT 59
US-08-980-400-2
; Sequence 2, Application US/08980400
; Patent No. 6018026
; GENERAL INFORMATION:
; APPLICANT: Sledziewski Ph.D., Andrzej Z
; APPLICANT: Bell, Lillian A.
; APPLICANT: Kindsvogel Ph.D., Wayne R.
; TITLE OF INVENTION: METHODS OF PRODUCING SECRETED RECEPTOR ANALOGS AND BIOLOGIC
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/980,400
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,329
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Marki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 990008.446C6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-622-4900
; TELEFAX: 206-682-6031
; TELEX: 3723836
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1106 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-980-400-2

Query Match
Best Local Similarity 40.0%; Score 6; DB 3; Length 1106;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGP 6
Db 656 SHLGP 661

RESULT 60
US-08-462-728-4
Sequence 4, Application US/08462728
Patent No. 6043211
GENERAL INFORMATION:
APPLICANT: WILLIAMS, Lewis T.
APPLICANT: ESCOBEDO, Jaime A.
TITLE OF INVENTION: PLATELET-DERIVED GROWTH FACTOR RECEPTORS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew
STREET: One Market, Steuart Street Tower, 20th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,728
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/309,322
FILING DATE: 10-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/151,414
FILING DATE: 02-FEB-1988
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 2307K-267-2-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/326-2400
TELEFAX: 415/326-2422
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1106 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-462-728-4

Query Match 40.0%; Score 6; DB 3; Length 1106;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGP 6
DB 656 SHLGP 661

RESULT 61
US-09-583-459A-2
Sequence 2, Application US/09583459A
Patent No. 6291212
GENERAL INFORMATION:
APPLICANT: Sledziewski Ph.D., Andrzej Z
APPLICANT: Bell, Lillian A.
APPLICANT: Kindsvogel Ph.D., Wayne R.
TITLE OF INVENTION: METHODS OF PRODUCING SECRETED RECEPTOR ANALOGS
AND BIOLOGICALLY ACTIVE DIMERIZED POLYPEPTIDE
TITLE OF INVENTION: FUSIONS
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry
STREET: 6300 Columbia Center
CITY: Seattle

STATE: WA
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/583,459A
FILING DATE: 30-MAY-2000
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/634,510
FILING DATE: 27-DEC-1990
APPLICATION NUMBER: US 07/146,877
FILING DATE: 22-JAN-1988
APPLICATION NUMBER: US 07/347,291
FILING DATE: 02-MAY-1989
ATTORNEY/AGENT INFORMATION:
NAME: Maki J.D., David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 990008.446C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031
TELEX: 3723836
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1106 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-583-459A-2

Query Match 40.0%; Score 6; DB 4; Length 1106;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGP 6
DB 656 SHLGP 661

RESULT 62
US-09-583-210-2
Sequence 2, Application US/09583210
Patent No. 6291646
GENERAL INFORMATION:
APPLICANT: Sledziewski Ph.D., Andrzej Z
APPLICANT: Bell, Lillian A.
APPLICANT: Kindsvogel Ph.D., Wayne R.
TITLE OF INVENTION: METHODS OF PRODUCING SECRETED RECEPTOR ANALOGS
AND BIOLOGICALLY ACTIVE DIMERIZED POLYPEPTIDE
TITLE OF INVENTION: FUSIONS
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry
STREET: 6300 Columbia Center
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/583,210
FILING DATE: 30-MAY-2000
CLASSIFICATION:
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/634,510
FILING DATE: 27-DEC-1990
APPLICATION NUMBER: US 07/146,877
FILING DATE: 22-JAN-1988
APPLICATION NUMBER: US 07/347,291
FILING DATE: 02-MAY-1989
ATTORNEY/AGENT INFORMATION:
NAME: MAKI J.D., David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 990008.446C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031
TELEX: 3723836
INFORMATION FOR SEQ. ID NO.: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1106 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-583-210-2

Query Match 40.0%; Score 6; DB 4; Length 1106;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGP 6
DB 656 SHLGP 661

RESULT 63
US-09-583-449A-2
Sequence 2, Application US/09583449A
Patent No. 630099
GENERAL INFORMATION:
APPLICANT: Sledziewski Ph.D., Andrzej Z
APPLICANT: Bell, Lillian A.
APPLICANT: Kindsvogel Ph.D., Wayne R.
TITLE OF INVENTION: METHODS OF PRODUCING SECRETED RECEPTOR ANALOGS
TITLE OF INVENTION: AND BIOLOGICALLY ACTIVE DIMERIZED POLYPEPTIDE
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry
STREET: 6300 Columbia Center
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/583,449A
FILING DATE: 30-MAY-2000
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/634,510
FILING DATE: 27-DEC-1990
APPLICATION NUMBER: US 07/146,877
FILING DATE: 22-JAN-1988
APPLICATION NUMBER: US 07/347,291
FILING DATE: 02-MAY-1989
ATTORNEY/AGENT INFORMATION:
NAME: MAKI J.D., David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 990008.446C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031

TELEX: 3723836
INFORMATION FOR SEQ. ID NO.: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1106 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-583-449A-2

Query Match 40.0%; Score 6; DB 4; Length 1106;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGP 6
DB 656 SHLGP 661

RESULT 64
US-09-435-059-2
Sequence 2, Application US/09435059
Patent No. 632323
GENERAL INFORMATION:
APPLICANT: Sledziewski Ph.D., Andrzej Z
APPLICANT: Bell, Lillian A.
APPLICANT: Kindsvogel Ph.D., Wayne R.
TITLE OF INVENTION: METHODS OF PRODUCING SECRETED RECEPTOR ANALOGS AND BIOLOGI
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/435,059
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/477,329
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: MAKI, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 990008.446C6
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031
TELEX: 3723836
INFORMATION FOR SEQ. ID NO.: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1106 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-435-059-2

Query Match 40.0%; Score 6; DB 4; Length 1106;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGP 6
DB 656 SHLGP 661

RESULT 65

US-08-461-917-4
; Sequence 4, Application US/08461917
; Patent No. 6372438
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, Lewis T.
; APPLICANT: ESCOBEDO, Jaime A.
; TITLE OF INVENTION: PLATELET-DERIVED GROWTH FACTOR RECEPTORS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: One Market, Stewart Street Tower, 20th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,917
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/309,322
; FILING DATE: 10-FEB-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/151,414
; FILING DATE: 02-FEB-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 2307K-267-2-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/326-2400
; TELEFAX: 415/326-2422
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1106 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-461-917-4

Query Match 40.0%; Score 6; DB 4; Length 1106;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGP 6
DB 656 SHLGP 661

RESULT 66
PCT-US92-00730-2
; Sequence 2, Application PC/TUS9200730
; GENERAL INFORMATION:
; APPLICANT: Wolf, David
; APPLICANT: Tomlinson, James E.
; APPLICANT: Fretto, Larry J.
; APPLICANT: Giese, Neill A.
; APPLICANT: Escobedo, Jaime A.
; APPLICANT: Williams, Lewis T.
; TITLE OF INVENTION: DOMAINS OF EXTRACELLULAR REGION OF HUMAN
; TITLE OF INVENTION: PLATELET-DERIVED GROWTH FACTOR RECEPTOR POLYPEPTIDES
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND
; STREET: Stewart Street Tower, 20th Floor \ One Market
; CITY: San Francisco
; STATE: California

COUNTRY: US
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/00730
; FILING DATE: 19920128
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: 12418-14
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1106 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US92-00730-2

Query Match 40.0%; Score 6; DB 5; Length 1106;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGP 6
DB 656 SHLGP 661

RESULT 67
PCT-US92-00862-2
; Sequence 2, Application PC/TUS9200862
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, Lewis T.
; APPLICANT: ESCOBEDO, Jaime A.
; TITLE OF INVENTION: PLATELET-DERIVED GROWTH FACTOR RECEPTORS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND
; STREET: Stewart Street Tower, 20th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/00862
; FILING DATE: 19920131
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/151,141
; FILING DATE: 02-FEB-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/309,322
; FILING DATE: 10-FEB-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: 2307U-267-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 1106 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US92-00862-2

Query Match 40.0%; Score 6; DB 5; Length 1106;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGP 6
|||||
DB 656 SHLGP 661

RESULT 68
US-08-984-709A-50
Sequence 50, Application US/08984709A
Patent No. 6320032
GENERAL INFORMATION:
APPLICANT: Williams, Mark E.
APPLICANT: Stauderman, Kenneth A.
APPLICANT: Harpold, Michael M.
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
METHODS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, Suite 700
CITY: La Jolla
STATE: California
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSRO Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/984,709A
FILING DATE: 02-DEC-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24735-9815 (formerly 6362-9815)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 450-8400
TELEFAX: (619) 587-5360
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 2353 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: Internal
ORIGINAL SOURCE:
US-08-984-709A-50

Query Match 40.0%; Score 6; DB 4; Length 2353;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 TPESRA 14
|||||
DB 2269 TPESRA 2294

RESULT 69

US-08-804-227C-14
Sequence 14, Application US/08804227C
Patent No. 5876991
GENERAL INFORMATION:
APPLICANT: Dehoff, Bradley S.
APPLICANT: Kuhstoss, Stuart A.
APPLICANT: Rostock, Paul R., Jr.
APPLICANT: Sutton, Kimberly L.
TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: THOMAS G. PLANT 1501
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCI(DOS) Text only
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,227C
FILING DATE: February 21, 1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Plant, Thomas, G.
REGISTRATION NUMBER: 35,784
REFERENCE/DOCKET NUMBER: X-8231
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-2459
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 4545 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-804-227C-14

Query Match 40.0%; Score 6; DB 2; Length 4545;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRA 15
|||||
DB 4008 PESRA 4013

RESULT 70
US-08-804-227C-8
Sequence 8, Application US/08804227C
Patent No. 5876991
GENERAL INFORMATION:
APPLICANT: Dehoff, Bradley S.
APPLICANT: Kuhstoss, Stuart A.
APPLICANT: Rostock, Paul R., Jr.
APPLICANT: Sutton, Kimberly L.
TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: THOMAS G. PLANT 1501
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCI(DOS) Text only
CURRENT APPLICATION DATA:

```

; APPLICATION NUMBER: US/08/804,227C
; FILING DATE: February 21, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Plant, Thomas, G.
; REGISTRATION NUMBER: 35,784
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-2459
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4550 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-804-227C-8

Query Match
Best Local Similarity 100.0%; Score 6; DB 2; Length 4550;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRAA 15
Db 4013 PESRAA 4018

RESULT 71
US-08-804-198-2
; Sequence 2, Application US/08804198
; Patent No. 5945320
; GENERAL INFORMATION:
; APPLICANT: Burgett, Stanley G.
; APPLICANT: Kunstoss, Stuart A.
; APPLICANT: Rao, Nagaraja R.
; APPLICANT: Richardson, Mark A.
; APPLICANT: Rosteck, Paul R., Jr.
; TITLE OF INVENTION: PLATENOLOIDE SYNTHASE GENE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PAUL R. CANTRELL 1138
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,198
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CANTRELL, PAUL R.
; REGISTRATION NUMBER: 36,470
; REFERENCE/DOCKET NUMBER: P9113
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3885
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4550 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-804-198-2

Query Match
Best Local Similarity 100.0%; Score 6; DB 2; Length 4550;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRAA 15
```

```

Db 4013 PESRAA 4018

RESULT 72
US-08-707-873-1
; Sequence 1, Application US/08707873
; Patent No. 5747318
; GENERAL INFORMATION:
; APPLICANT: KARIN, MICHAEL
; APPLICANT: DENG, TILIANG
; TITLE OF INVENTION: NOVEL MITOGEN ACTIVATED PROTEIN KINASE,
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
; STREET: 1880 CENTURY PARK EAST, FIFTH FLOOR
; CITY: LOS ANGELES
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/707,873
; FILING DATE: 09-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/315,067
; FILING DATE: 29-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: HAILLE PH. D., LISA A.
; REGISTRATION NUMBER: P-38,347
; REFERENCE/DOCKET NUMBER: PD3743
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/455-5100
; TELEFAX: 619/455-5110
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..11
US-08-707-873-1

Query Match
Best Local Similarity 100.0%; Score 5; DB 1; Length 11;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPES 12
Db 6 STPES 10

RESULT 73
US-08-707-874-1
; Sequence 1, Application US/08707874
; Patent No. 5817451
; GENERAL INFORMATION:
; APPLICANT: KARIN, MICHAEL
; APPLICANT: DENG, TILIANG
; TITLE OF INVENTION: NOVEL MITOGEN ACTIVATED PROTEIN KINASE,
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
```

STREET: 1880 CENTURY PARK EAST, FIFTH FLOOR
CITY: LOS ANGELES
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/707,874
FILING DATE: 09-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/315,067
FILING DATE: 29-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: HAILE PH.D., LISA A.
REGISTRATION NUMBER: P-38,347
REFERENCE/DOCKET NUMBER: PD3743
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/455-5100
TELEFAX: 619/455-5110
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..11
US-08-707-874-1

Query Match 33.3%; Score 5; DB 2; Length 11;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STEPS 12
Db 6 STEPS 10

RESULT 74
US-08-315-067-1
Sequence 1, Application US/08315067
Patent No. 5925557
GENERAL INFORMATION:
APPLICANT: KARIN, MICHAEL
APPLICANT: DENG, TILIANG
TITLE OF INVENTION: NOVEL MITOGEN ACTIVATED PROTEIN KINASE,
TITLE OF INVENTION: FRK
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
STREET: 1880 CENTURY PARK EAST, FIFTH FLOOR
CITY: LOS ANGELES
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/315,067
FILING DATE: 29-SEP-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: HAILE PH.D., LISA A.

REGISTRATION NUMBER: P-38,347
REFERENCE/DOCKET NUMBER: PD3743
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/455-5100
TELEFAX: 619/455-5110
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..11
US-08-315-067-1

Query Match 33.3%; Score 5; DB 2; Length 11;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STEPS 12
Db 6 STEPS 10

RESULT 75
US-09-193-797-1
Sequence 1, Application US/09193797
Patent No. 6054560
GENERAL INFORMATION:
APPLICANT: KARIN, MICHAEL
APPLICANT: DENG, TILIANG
TITLE OF INVENTION: NOVEL MITOGEN ACTIVATED PROTEIN KINASE,
TITLE OF INVENTION: FRK
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
STREET: 1880 CENTURY PARK EAST, FIFTH FLOOR
CITY: LOS ANGELES
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/193,797
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/315,067
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: HAILE PH.D., LISA A.
REGISTRATION NUMBER: P-38,347
REFERENCE/DOCKET NUMBER: PD3743
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/455-5100
TELEFAX: 619/455-5110
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..11
US-09-193-797-1

Query Match 33.3%; Score 5; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STEPS 12
|||||
DB 6 STEPS 10

RESULT 76
PCT-US95-01770-1
Sequence 1, Application PC/TUS9501770
GENERAL INFORMATION:
APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
TITLE OF INVENTION: NOVEL MITOGEN ACTIVATED PROTEIN KINASE,
TITLE OF INVENTION: FRK
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: ROBBINS, BERLINER & CARSON
STREET: 201 N. FIGUEROA STREET, 5TH FLOOR
CITY: LOS ANGELES
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 90012-2628
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/01770
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: BERLINER, ROBERT
REGISTRATION NUMBER: 20,121
REFERENCE/DOCKET NUMBER: 5555-297
TELECOMMUNICATION INFORMATION:
TELEPHONE: 213-977-1001
TELEFAX: 213-977-1003
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..11
PCT-US95-01770-1

Query Match 33.3%; Score 5; DB 5; Length 11;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STEPS 12
|||||
DB 6 STEPS 10

RESULT 77
US-08-630-897-1
Sequence 1, Application US/08630897
Patent No. 5639601
GENERAL INFORMATION:
APPLICANT: SAKETI, TAKAKIYO
TITLE OF INVENTION: PEPTIDE FOR DIAGNOSIS OF FOOT-AND-MOUTH
DISEASE, AND ANTIGENS CONTAINING THE PEPTIDE FOR DIAGNOSIS
TITLE OF INVENTION: OF FOOT-AND-MOUTH DISEASE
NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: USA
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,897
FILING DATE: 04-APR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 148364/1995
FILING DATE: 24-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 7614-001-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-630-897-1

Query Match 33.3%; Score 5; DB 1; Length 13;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPE 11
|||||
DB 1 RSTPE 5

RESULT 78
PCT-US93-06751-4
Sequence 4, Application PC/TUS9306751
GENERAL INFORMATION:
APPLICANT: P. Keller, A. J. Conley, A.R. Shaw, B.A. Arnold
TITLE OF INVENTION: Immunological conjugates of OMPC and
TITLE OF INVENTION: HIV-Specific Selected Principal Neutralization GXG Epitopes
NUMBER OF SEQUENCES: 146
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/06751
FILING DATE: 19930719
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Meridith, Roy D.
REGISTRATION NUMBER: 30,777
REFERENCE/DOCKET NUMBER: 18614

TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-4678
TELEFAX: (908) 594-4720
TELEX: 138825
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE: Random Epitope Library Alpha
PCT-US93-06751-4

Query Match 33.3%; Score 5; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 GPHRS 8
Db 11 GPHRS 15

RESULT 79
US-08-751-767A-67
Sequence 67, Application US/08751767A
Patent No. 5994104
GENERAL INFORMATION:
APPLICANT: ANDERSON, ROBERT J.
APPLICANT: GRANT, HUGH
APPLICANT: MACDONALD, IAN D.
TITLE OF INVENTION: INTERLUKIN-12 FUSION PROTEIN
NUMBER OF SEQUENCES: 80
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/751,767A
FILING DATE: 08-NOV-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 117-221
TELECOMMUNICATION INFORMATION:
TELEPHONE: 7038164091
TELEFAX: 7038164100
INFORMATION FOR SEQ ID NO: 67:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-751-767A-67

Query Match 33.3%; Score 5; DB 2; Length 26;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 GPHRS 8
Db 3 GPHRS 7

RESULT 80
US-07-743-518-18
Sequence 18, Application US/07743518
Patent No. 5397696
GENERAL INFORMATION:
APPLICANT: YANAGIHARA, RICHARD
APPLICANT: NERURKAR, VIVEK R.
APPLICANT: JENKINS, CAROL
APPLICANT: MILLER, MARK
APPLICANT: GARRUTO, RALPH M.
TITLE OF INVENTION: PAPUA NEW GUINEA HUMAN T-LYMPHOTROPIC
VIRUS
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
STREET: 1615 L STREET, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20036-5601

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/743,518
FILING DATE: 19910812
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: SCOTT, MATSON T.
REGISTRATION NUMBER: 26,581
REFERENCE/DOCKET NUMBER: WTS/5683/84699/SAP

TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-743-518-18

Query Match 33.3%; Score 5; DB 1; Length 40;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HLGPH 6
Db 36 HLGPH 40

RESULT 81
US-08-466-886-41
Sequence 41, Application US/08466886
Patent No. 5776677
GENERAL INFORMATION:
APPLICANT: TSUI, LAP-Chee
APPLICANT: Riordan, John R.
APPLICANT: Rommens, Johanna M.
APPLICANT: Kerem, Bat-Sheva
APPLICANT: Collins, Francis S.
APPLICANT: Iannuzzi, Michael C.
APPLICANT: Drumm, Mitchell L.
APPLICANT: Buckwald, Manuel
TITLE OF INVENTION: Cystic Fibrosis Gene
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX

STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,886
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Goldstein, Jorge A.
REGISTRATION NUMBER: 29,021
REFERENCE/DOCKET NUMBER: 1329.0010006
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 111 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-466-886-41

Query Match 33.3%; Score 5; DB 1; Length 111;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRA 14
|||||
DB 58 PESRA 62

RESULT 82
US-08-469-617-41
Sequence 41, Application US/08469617
Patent No. 6201107
GENERAL INFORMATION:
APPLICANT: Tsui, Lap-Chee
APPLICANT: Riordan, John R.
APPLICANT: Rommens, Johanna M.
APPLICANT: Kerem, Bat-Sheva
APPLICANT: Collins, Francis S.
APPLICANT: Iannuzzi, Michael C.
APPLICANT: Drumm, Mitchell L.
APPLICANT: Buckwald, Manuel
TITLE OF INVENTION: Cystic Fibrosis Gene
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,617
FILING DATE: 06-JUN-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Goldstein, Jorge A.
REGISTRATION NUMBER: 29,021

REFERENCE/DOCKET NUMBER: 1329.0010008
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 111 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-469-617-41

Query Match 33.3%; Score 5; DB 4; Length 111;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRA 14
|||||
DB 58 PESRA 62

RESULT 83
US-09-134-001C-4265
Sequence 4265, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4265
LENGTH: 161
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4265

Query Match 33.3%; Score 5; DB 4; Length 161;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGP 5
|||||
DB 99 SHLGP 103

RESULT 84
US-09-069-896-1
Sequence 1, Application US/09069896
Patent No. 6071720
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Patterson, Chandra
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: DELAYED RECTIFIER POTASSIUM
TITLE OF INVENTION: CHANNEL HOMOLOG
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette

COMPUTER: IBM Compatidble
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/069,896
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cerrone, Michael C
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0507 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 170 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRSTNOT03
CLONE: 637471
US-09-069-896-1

Query Match 33.3%; Score 5; DB 3; Length 170;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRAA 15
Db 22 ESRAA 26

RESULT 85
US-09-471-468-1
Sequence 1, Application US/09471468
Patent No. 6432687
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Patterson, Chandra
TITLE OF INVENTION: DELAYED RECTIFIER POTASSIUM
TITLE OF INVENTION: CHANNEL HOMOLOG
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatidble
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/471,468
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/069,896
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cerrone, Michael C
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0507 US
TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 170 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRSTNOT03
CLONE: 637471
US-09-471-468-1

Query Match 33.3%; Score 5; DB 4; Length 170;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRAA 15
Db 22 ESRAA 26

RESULT 86
US-09-413-814-98
Sequence 98, Application US/09413814
Patent No. 6225064
GENERAL INFORMATION:
APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
APPLICANT: Bristol-Myers Squibb, Co.
APPLICANT: Beyer, Stefan
APPLICANT: Bloeker, Helmut
APPLICANT: Brandt, Petra
APPLICANT: Cino, Paul M
APPLICANT: Dougherty, Brian A
APPLICANT: Goldberg, Steven L
APPLICANT: Hoffer, Gerhard
APPLICANT: Mueller, Joachim
TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polypeptide or
TITLE OF INVENTION: heteropolypeptide compounds
FILE REFERENCE: PCT/US 99/23535
CURRENT APPLICATION NUMBER: US/09/413,814
CURRENT FILING DATE: 1999-10-07
EARLIER APPLICATION NUMBER: DE 198 46 493.2
EARLIER FILING DATE: 1998-10-09
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 98
LENGTH: 242
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-413-814-98

Query Match 33.3%; Score 5; DB 4; Length 242;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRAA 15
Db 154 ESRAA 158

RESULT 87
US-09-530-058-6
Sequence 6, Application US/09530058
Patent No. 6379938
GENERAL INFORMATION:
APPLICANT: <Unknown>
TITLE OF INVENTION: Epoxide hydrolase
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear

STREET: 501 West Broadway
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-3505
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Win95
SOFTWARE: Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/530,058
FILING DATE: 03-Jul-2000
ATTORNEY/AGENT INFORMATION:
NAME: Hart, Daniel
REGISTRATION NUMBER: 40,637
REFERENCE/DOCKET NUMBER: VANN150.001APC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 253 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-530-058-6
Query Match 33.3%; Score 5; DB 4; Length 253;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5 PRST 9
DB 165 PRST 169
RESULT 88
US-08-911-853-37
Sequence 37, Application US/08911853
Patent No. 6048710
GENERAL INFORMATION:
APPLICANT: Gerltse, Gjsbert
ATTORNEY/AGENT INFORMATION:
NAME: Quax, Wilhelmus J.
REGISTRATION NUMBER: 37
TITLE OF INVENTION: EXPRESSION SYSTEM FOR ALTERED
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESS: Genencor International
STREET: 925 Page Mill Road
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1013
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/911,853
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/699,092
FILING DATE: 16-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gialster, Debra J
REGISTRATION NUMBER: 33,888
REFERENCE/DOCKET NUMBER: GC361-2

TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-846-7620
TELEFAX: 650-845-6504
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 282 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-911-853-37

Query Match 33.3%; Score 5; DB 3; Length 282;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTP 11
DB 20 RSTP 24

RESULT 89
US-09-479-409-37
Sequence 37, Application US/09479409
Patent No. 6225106
GENERAL INFORMATION:
APPLICANT: Gerltse, Gjsbert
ATTORNEY/AGENT INFORMATION:
NAME: Quax, Wilhelmus J.
REGISTRATION NUMBER: 37
TITLE OF INVENTION: EXPRESSION SYSTEM FOR ALTERED
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESS: Genencor International
STREET: 925 Page Mill Road
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1013
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/479,409
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/911,853
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Gialster, Debra J
REGISTRATION NUMBER: 33,888
REFERENCE/DOCKET NUMBER: GC361-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-846-7620
TELEFAX: 650-845-6504
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 282 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-479-409-37
Query Match 33.3%; Score 5; DB 4; Length 282;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 RSTP 11
DB 20 RSTP 24
RESULT 90

US-09-479-453-37
; Sequence 37, Application US/09479453
; Patent No. 6313283
; GENERAL INFORMATION:
; APPLICANT: Gerltse, Gijstbert
; APPLICANT: Quax, Wilhelmus J.
; TITLE OF INVENTION: EXPRESSION SYSTEM FOR ALTERED
; TITLE OF INVENTION: EXPRESSION LEVELS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genencor International
; STREET: 925 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1013
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/479,453
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/911,853
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Gjalster, Debra J
; REGISTRATION NUMBER: 33, 888
; REFERENCE/DOCKET NUMBER: GC361-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-845-7620
; TELEFAX: 650-845-6504
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 282 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-479-453-37
Query Match 33.3%; Score 5; DB 4; Length 282;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 RSTPE 11
DB 20 RSTPE 24
RESULT 91
US-08-701-191A-16
; Sequence 16, Application US/08701191A
; Patent No. 5942428
; GENERAL INFORMATION:
; APPLICANT: Moosa Mohammadi, Joseph Schlessinger,
; APPLICANT: and Stevan R. Hubbard
; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN
; TITLE OF INVENTION: OF NON-INSULIN RECEPTOR TYROSINE KINASE
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 MB
; COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/701,191A
; FILING DATE: August 21, 1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 227/088
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 293 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-701-191A-16
Query Match 33.3%; Score 5; DB 2; Length 293;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HLGPH 6
DB 71 HLGPH 75
RESULT 92
US-09-323-872A-34
; Sequence 34, Application US/09323872A
; Patent No. 6395539
; GENERAL INFORMATION:
; APPLICANT: Coschigano, Peter
; TITLE OF INVENTION: Compositions and Methods for Bioremediation
; FILE REFERENCE: OHU-03640
; CURRENT APPLICATION NUMBER: US/09/323,872A
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 09/072,433
; PRIOR FILING DATE: 1998-05-04
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 34
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-09-323-872A-34
Query Match 33.3%; Score 5; DB 4; Length 308;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 10 PESRA 14
DB 44 PESRA 48
RESULT 93
US-09-199-637A-23
; Sequence 23, Application US/09199637A
; Patent No. 6355411
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick
; APPLICANT: Goodman, Howard M.
; APPLICANT: Rahme, Laurence G.
; APPLICANT: Mahajan-Miklos, Shalima

APPLICANT: Tan, Man-Wah
APPLICANT: Cao, Hui
APPLICANT: Drenkard, Eliana
APPLICANT: Tsongalis, John
TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
TITLE OF INVENTION: SEQUENCES AND USES THEREOF
FILE REFERENCE: 00786/361002
CURRENT APPLICATION NUMBER: US/09/199,637A
CURRENT FILING DATE: 1998-11-25
PRIOR APPLICATION NUMBER: 60/066,517
PRIOR FILING DATE: 1997-11-25
NUMBER OF SEQ ID NOS: 437
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 23
LENGTH: 318
TYPE: PR
ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-23

Query Match 33.3%; Score 5; DB 4; Length 318;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 HRSTP 10
DB 13 HRSTP 17

RESULT 94
US-08-712-948-2
Sequence 2, Application US/08712948
Patent No. 5850002
GENERAL INFORMATION:
APPLICANT: Korsmeyer, Stanley J.
TITLE OF INVENTION: HOX11 Gain and Loss of Function Murine
TITLE OF INVENTION: Models
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 1100 Peachtree Street, Suite 2800
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-4530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/712,948
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/231,728
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: W0104
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-815-6508
TELEFAX: (404)-815-6555
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 330 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal

ORIGINAL SOURCE:
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc_feature
LOCATION: 200..260
OTHER INFORMATION: /function= "homeobox domain"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 19..25
OTHER INFORMATION: /function= "Hep motif"
PUBLICATION INFORMATION:
AUTHORS: Hatanou,
AUTHORS: Roberts,
AUTHORS: Minden,
AUTHORS: Crist,
AUTHORS: Korsmeyer,
TITLE: Deregulation of a Homeobox gene, HOX11, by
TITLE: the t(10;14) in T Cell Leukemia
JOURNAL: Science
VOLUME: 253
PAGES: 79-82
DATE: July 5-1991
RELEVANT RESIDUES IN SEQ ID NO: 2: FROM 1 TO 330
US-08-712-948-2

Query Match 33.3%; Score 5; DB 2; Length 330;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HLGPH 6
DB 3 HLGPH 7

RESULT 95
US-08-712-948-1
Sequence 1, Application US/08712948
Patent No. 5850002
GENERAL INFORMATION:
APPLICANT: Korsmeyer, Stanley J.
TITLE OF INVENTION: HOX11 Gain and Loss of Function Murine
TITLE OF INVENTION: Models
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 1100 Peachtree Street, Suite 2800
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-4530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/712,948
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/231,728
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: W0104
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-815-6508
TELEFAX: (404)-815-6555
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 333 amino acids
TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Murine
US-08-712-948-1

Query Match 33.3%; Score 5; DB 2; Length 333;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HUGPH 6
DB 3 HUGPH 7

RESULT 96
US-09-453-702B-263
Sequence 263, Application US/09453702B
Patent No. 6365723
GENERAL INFORMATION:
APPLICANT: Blatner, Frederick R.
Burland, Valerie
Perna, Nicole T.
Plunkett, Guy
Welch, Rod

TITLE OF INVENTION: No. 6365723el Sequences of E. coli 0157
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Charles & Brady
STREET: 1 South Pluckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53701-2113

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 8.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/453,702B
FILING DATE: 03-Dec-1999
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/110,955
FILING DATE: 04-Dec-1998
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296, 95017

TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 263:
SEQUENCE CHARACTERISTICS:
LENGTH: 348 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>

MOLECULE TYPE: protein
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 263:
US-09-453-702B-263

Query Match 33.3%; Score 5; DB 4; Length 348;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 11 ESRAA 15
|||||

DB 160 ESRAA 164

RESULT 97
US-09-302-769-25
Sequence 25, Application US/09302769
Patent No. 6323317
GENERAL INFORMATION:
APPLICANT: HILTON, Douglas J
APPLICANT: ALEXANDER, Warren S
APPLICANT: VINEY, Elizabeth M
APPLICANT: WILSON, Tracey A
APPLICANT: RICHARDSON, Rachael T
APPLICANT: STARK, Robyn
APPLICANT: NICHOLSON, Sandra E
APPLICANT: METCALF, Donald
APPLICANT: NICOLA, Nicos A

TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC AGENTS
FILE REFERENCE: 109762
CURRENT APPLICATION NUMBER: US/09/302,769
CURRENT FILING DATE: 1999-04-30
PRIOR APPLICATION NUMBER: 08/962,560
PRIOR FILING DATE: 1997-10-31

NUMBER OF SEQ ID NOS: 50
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 25
LENGTH: 350
TYPE: PRT
ORGANISM: Mouse

FEATURE:
NAME/KEY: UNSURE
LOCATION: (167)
OTHER INFORMATION: Xaa 1s unsure
US-09-302-769-25

Query Match 33.3%; Score 5; DB 4; Length 350;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 HRSTP 10
DB 158 HRSTP 162

RESULT 98
US-08-239-431A-4
Sequence 4, Application US/08239431A
Patent No. 5716835
GENERAL INFORMATION:
APPLICANT: Regan, John W.
APPLICANT: Gil, Daniel W.
APPLICANT: Woodward, David F.

TITLE OF INVENTION: NOVEL HUMAN EP PROSTAGLANDIN RECEPTOR
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson and Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: USA
ZIP: 92660

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/239,431A
FILING DATE: 05-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned A
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: ALNGN.053A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 358 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: Internal
ORIGINAL SOURCE:
US-08-239-431A-4

Query Match
Best Local Similarity 33.3%; Score 5; DB 1; Length 358;
100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRA 14
11111
Db 104 PESRA 108

RESULT 99
US-08-463-081B-6
Sequence 6, Application US/08463081B
Patent No. 5871960
Patent No. 5871960 5837487
GENERAL INFORMATION:
APPLICANT: Smith, Kendall A. & Beadling, Carol
TITLE OF INVENTION: Nucleic Acids Encoding CR5 Polypeptide,
TITLE OF INVENTION: Vector and Transformed Cell Thereof, and Expression Thereof
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: PRETTY, SCHROEDER & POPLANSKI
STREET: 444 South Flower St. - Suite 1900
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0,
SOFTWARE: Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,081B
FILING DATE: 5-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/104,736
FILING DATE: 10-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/796,066
FILING DATE: 20-NOV-91
ATTORNEY/AGENT INFORMATION:
NAME: Viviana Amzel, Ph. D.
REGISTRATION NUMBER: 30,930
REFERENCE/DOCKET NUMBER: P66 38150 (DART-060)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 622-7700
TELEFAX: (213) 489-4210
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 358 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-463-081B-6

Query Match
Best Local Similarity 33.3%; Score 5; DB 2; Length 358;
100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRA 14
11111
Db 104 PESRA 108

RESULT 100
US-08-461-379A-6
Sequence 6, Application US/08461379A
Patent No. 5871961
GENERAL INFORMATION:
APPLICANT: Smith, Kendall A. & Beadling, Carol
TITLE OF INVENTION: Nucleic Acids Encoding CR5 Polypeptide,
TITLE OF INVENTION: Vector and Transformed Cell Thereof, and
TITLE OF INVENTION: Expression Thereof
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ratner & Prestia
ADDRESS: (B) STREET: One Westlakes-Beryn
CITY: Valley Forge
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0,
SOFTWARE: Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,379A
FILING DATE: 5-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/330,108; 08/104,736
APPLICATION NUMBER: 807/796,066
FILING DATE: 27-OCT-1994; 10-AUG-1993 & 20-NOV-91
ATTORNEY/AGENT INFORMATION:
NAME: Viviana Amzel, Ph. D.
REGISTRATION NUMBER: 30,930
REFERENCE/DOCKET NUMBER: DART-070
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610)470-0700
TELEFAX: (610)470-0701
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 358 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-461-379A-6

Query Match
Best Local Similarity 33.3%; Score 5; DB 2; Length 358;
100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRA 14
11111
Db 104 PESRA 108

RESULT 101
US-08-462-390B-6
Sequence 6, Application US/08462390B
Patent No. 5882894
GENERAL INFORMATION:
APPLICANT: Smith, K. A., & Beadling, C.

;; TITLE OF INVENTION: Nucleic Acids Encoding CR8 Polypeptide, Vector and
;; TITLE OF INVENTION: Transformed Cell Thereof, and Expression Thereof
;; NUMBER OF SEQUENCES: 35
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Ratner & Prestia
;; (B) STREET: One Westlakes-Berwyn
;; CITY: Valley Forge
;; STATE: Pennsylvania
;; COUNTRY: USA
;; ZIP: 19482
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/462,390B
;; FILING DATE: 5-JUNE-1995
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: USSN 08/330,108
;; FILING DATE: 27-OCT-1994
;; APPLICATION NUMBER: USSN 08/104,736
;; FILING DATE: 10-AUG-1993
;; APPLICATION NUMBER: USSN 07/796,066
;; FILING DATE: 20-NOV-91
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Viviana Amzel, Ph. D.
;; REGISTRATION NUMBER: 30,930
;; REFERENCE/DOCKET NUMBER: DART-040
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (610)407-0700
;; TELEFAX: (610)407-0701
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 358 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-462-390B-6
;
Query Match 33.3%; Score 5; DB 2; Length 358;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
QY 10 PESRA 14
DB 104 PESRA 108
;
RESULT 102
US-08-463-074B-6
;; Sequence 6, Application US/08463074B
;; Patent No. 6020155
;; GENERAL INFORMATION:
;; APPLICANT: Smith, Kendall A. & Beadling, Carol
;; TITLE OF INVENTION: Nucleic Acids Encoding CRI Fusion protein, Vector an
;; NUMBER OF SEQUENCES: 35
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: PRETTY, SCHROEDER & POPLAWSKI
;; (B) STREET:
;; CITY: Los Angeles
;; STATE: California
;; COUNTRY: USA
;; ZIP: 90071
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0,
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/463,074B
;; FILING DATE: 5-JUN-1995

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/104,736
;; FILING DATE: 10-AUG-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/796,066
;; FILING DATE: 20-NOV-91
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Viviana Amzel, Ph. D.
;; REGISTRATION NUMBER: 30,930
;; REFERENCE/DOCKET NUMBER: P66 38143 (DART-020)
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (213) 622-7700
;; TELEFAX: (213) 489-4210
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 358 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-463-074B-6
;
Query Match 33.3%; Score 5; DB 3; Length 358;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
QY 10 PESRA 14
DB 104 PESRA 108
;
RESULT 103
US-08-465-585C-6
;; Sequence 6, Application US/08465585C
;; Patent No. 6027914
;; GENERAL INFORMATION:
;; APPLICANT: Smith, K. A. & Beadling, C.
;; TITLE OF INVENTION: Nucleic Acids Encoding CR6 Polypeptide, Vecto
;; NUMBER OF SEQUENCES: 35
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: PRETTY, SCHROEDER & POPLAWSKI
;; (B) STREET:
;; CITY: Los Angeles
;; STATE: California
;; COUNTRY: USA
;; ZIP: 900071
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/465,585C
;; FILING DATE: 5-JUNE-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: USSN 08/330,108
;; FILING DATE: 27-OCT-1994
;; APPLICATION NUMBER: USSN 08/104,736
;; FILING DATE: 10-AUG-1993
;; APPLICATION NUMBER: USSN 07/796,066
;; FILING DATE: 20-NOV-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Viviana Amzel, Ph. D.
;; REGISTRATION NUMBER: 30,930
;; REFERENCE/DOCKET NUMBER: P66 38149 (DART-050)
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (213) 622-7700
;; TELEFAX: (213) 4894210
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 358 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-465-585C-6
Query Match 33.3%; Score 5; DB 3; Length 358;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 10 PESRA 14
DB 104 PESRA 108
RESULT 104
US-08-652-446-6
Sequence 6, Application US/08652446
Patent No. 6057427
GENERAL INFORMATION:
APPLICANT: Smith, Kendall A. & Beadling, Carol
TITLE OF INVENTION: Nucleic Acids Encoding CRS
TITLE OF INVENTION: Polypeptide, Vector and Transformed Cell Thereof, and
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: PRETTY, SCHROEDER & POPLANSKI
ADDRESS: (B) STREET:
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
444 South Flower St. - Suite 1900
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0,
SOFTWARE: Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/652,446
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP App. # 96921319.8
FILING DATE: 5-JAN-1998
APPLICATION NUMBER: PCT/US/96/09194
FILING DATE: 5-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/330,108
FILING DATE: 27-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/463,074
FILING DATE: 5-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,337
FILING DATE: 5-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,390
FILING DATE: 5-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/465,585
FILING DATE: 5-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/463,081
FILING DATE: 5-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/461,379
FILING DATE: 5-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/739,523
FILING DATE: 29-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Viviana Amzel, Ph. D.
REGISTRATION NUMBER: 30,930
REFERENCE/DOCKET NUMBER: PP66 40035
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 622-7700

TELEFAX: (213) 489-4210
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 358 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-652-446-6
Query Match 33.3%; Score 5; DB 3; Length 358;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 10 PESRA 14
DB 104 PESRA 108
RESULT 105
US-09-267-423-4
Sequence 4, Application US/09267423
Patent No. 6395878
GENERAL INFORMATION:
APPLICANT: Regan, John W.
APPLICANT: Gil, Daniel W.
APPLICANT: Woodward, David F.
TITLE OF INVENTION: No. 6395878el Human Prostaglandin EP Receptor
FILE REFERENCE: 17023 DIV CIP
CURRENT APPLICATION NUMBER: US/09/267,423
CURRENT FILING DATE: 1999-03-12
EARLIER APPLICATION NUMBER: 09/019,393
EARLIER FILING DATE: 1998-02-05
EARLIER APPLICATION NUMBER: 08/239,431
EARLIER FILING DATE: 1994-05-05
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 358
TYPE: PRT
ORGANISM: Homo sapiens
US-09-267-423-4
Query Match 33.3%; Score 5; DB 4; Length 358;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 10 PESRA 14
DB 104 PESRA 108
RESULT 106
US-09-552-322-2
Sequence 2, Application US/09552322
Patent No. 6436642
GENERAL INFORMATION:
APPLICANT: Gould-Rothberg
APPLICANT: Rastelli
TITLE OF INVENTION: METHOD OF CLASSIFYING A THYROID CARCINOMA USING
FILE REFERENCE: 15966-548
CURRENT APPLICATION NUMBER: US/09/552,322
CURRENT FILING DATE: 2000-04-19
PRIOR APPLICATION NUMBER: 60/130,123
PRIOR FILING DATE: 1999-04-20
PRIOR APPLICATION NUMBER: 60/193,203
PRIOR FILING DATE: 2000-03-30
NUMBER OF SEQ ID NOS: 63
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2
LENGTH: 384
TYPE: PRT
ORGANISM: Homo sapiens

US-09-552-322-2

Query Match 33.3%; Score 5; DB 4; Length 386;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0;

OY 10 PESRA 14
|||||
Db 274 PESRA 278

RESULT 107

US-09-390-326-6
; Sequence 6, Application US/09390326
; Patent No. 6316603
; GENERAL INFORMATION:
; APPLICANT: MCTIGUE, MICHELE A.
; APPLICANT: WICKERSHAM, JOHN A.
; APPLICANT: PINKO, CHRIS
; APPLICANT: SHOMALTER, RICHARD
; APPLICANT: PARAST, CAMRAN V.
; APPLICANT: TEMPCZYK-RUSSEL, ANNA
; APPLICANT: GEHRING, MICHAEL R.
; APPLICANT: MROCKZKOWSKI, BARBARA
; APPLICANT: KAN, CHEN-CHEN
; APPLICANT: VILLAFRANCA, J. ERNEST
; APPLICANT: APPELT, KRYSZTOF
; TITLE OF INVENTION: MODIFICATIONS OF THE VEGF RECEPTOR-2 PROTEIN AND
; FILE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 0125-001605
; CURRENT APPLICATION NUMBER: US/09/390,326
; CURRENT FILING DATE: 1999-09-07
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 386
; TYPE: PRT
; ORGANISM: E. coli
US-09-390-326-6

Query Match 33.3%; Score 5; DB 4; Length 386;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0;

OY 2 HLGPH 6
|||||
Db 75 HLGPH 79

RESULT 108

5171840-5
; Patent No. 5171840
; APPLICANT: KISHIMOTO, TADAMITSU
; TITLE OF INVENTION: RECEPTOR PROTEIN FOR HUMAN B CELL
; STIMULATORY FACTOR-2
; NUMBER OF SEQUENCES: 11
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/298,694
; FILING DATE: 19-JAN-1989
; SEQ ID NO: 5
; LENGTH: 386
5171840-5

Query Match 33.3%; Score 5; DB 6; Length 386;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0;

OY 7 RSTPE 11
|||||
Db 53 RSTPE 57

RESULT 109

5480796-5

; Patent No. 5480796
; APPLICANT: KISHIMOTO, TADAMITSU
; TITLE OF INVENTION: ANTIBODIES AGAINST THE RECEPTOR PROTEIN
; FOR HUMAN B CELL STIMULATORY FACTOR-2
; NUMBER OF SEQUENCES: 8
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/907,650
; FILING DATE: 02-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 298,694
; FILING DATE: 19-JAN-1989
; SEQ ID NO: 5
; LENGTH: 386
5480796-5

Query Match 33.3%; Score 5; DB 6; Length 386;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0;

OY 7 RSTPE 11
|||||
Db 53 RSTPE 57

RESULT 110

US-08-261-110A-4
; Sequence 4, Application US/08261110A
; Patent No. 5674992
; GENERAL INFORMATION:
; APPLICANT: JACENDORF, ANDRE
; APPLICANT: CERUTTI, HERIBERTO
; TITLE OF INVENTION: CDNA ENCODING A RECA HOMOLOG IN
; NUMBER OF SEQUENCES: 6
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP
; STREET: CLINTON SQUARE, P.O. BOX 1051
; CITY: ROCHESTER
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 14603

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/261,110A
; FILING DATE: 16-JUN-1994
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/938,332
; FILING DATE: 28-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: TIMMAN, SUSAN J.
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 19603/10231(D-1292A)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1636
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 388 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-261-110A-4

Query Match 33.3%; Score 5; DB 1; Length 388;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0;

QY 6 HRSTP 10
11111
DB 360 HRSTP 364

RESULT 111

5212296-9
; Patent No. 5212296
; APPLICANT: DEAN, CAROLINE; HARDER, PATRICIA A.; LEFTO, KENNETH
; J.; O'KEEFE, DANIEL P.; OMER, CHARLES A.; ROMESSER, JAMES A.
; PEPPERMAN, JAMES M.
; TITLE OF INVENTION: EXPRESSION OF HERBICIDE METABOLIZING
; CYTOCHROMES
; NUMBER OF SEQUENCES: 19
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/569,781
; FILING DATE: 23-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 464,499
; FILING DATE: 12-JAN-1990
; APPLICATION NUMBER: 405,605
; FILING DATE: 11-SEP-1989
; SEQ ID NO: 9:
5212296-9
; LENGTH: 403

Query Match 33.3%; Score 5; DB 6; Length 403;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPE 11
11111
DB 184 RSTPE 188

RESULT 112

US-08-725-758A-4
; Sequence 4, Application US/08725758A
; Patent No. 6160108
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy
; TITLE OF INVENTION: NOVEL PLATELET ACTIVATION PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/725,758A
; FILING DATE: 04-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/005,074
; FILING DATE: 06-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 05433/020001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 410 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: Internal
US-08-725-758A-4

Query Match 33.3%; Score 5; DB 4; Length 410;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STPEs 12
11111
DB 280 STPEs 284

RESULT 113

US-09-362-473-14
; Sequence 14, Application US/09362473
; Patent No. 6218169
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Edgar B.
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, S. Carl
; APPLICANT: Morgante, Michele
; APPLICANT: Rafalski, J. Antoni
; APPLICANT: Hitz, William D.
; APPLICANT: Kinney, Anthony J.
; TITLE OF INVENTION: Aromatic Amino Acid Catabolism Enzymes
; FILE REFERENCE: BB-1197
; CURRENT APPLICATION NUMBER: US/09/362,473
; CURRENT FILING DATE: 1999-07-28
; EARLIER APPLICATION NUMBER: 60/094,783
; EARLIER FILING DATE: JULY 31, 1998
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Glycine max
US-09-362-473-14

Query Match 33.3%; Score 5; DB 4; Length 421;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRA 14
11111
DB 361 PESRA 365

RESULT 114

US-08-725-758A-2
; Sequence 2, Application US/08725758A
; Patent No. 6160108
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy
; TITLE OF INVENTION: NOVEL PLATELET ACTIVATION PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 2.0
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/725,758A
FILING DATE: 04-09-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/005,074
FILING DATE: 06-09-1995
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 05433/020001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 426 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-725-758a-2

Query Match 33.3%; Score 5; DB 4; Length 426;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STEPS 12
Db 280 STEPS 284

RESULT 115
US-08-845-258-34
Sequence 34, Application US/08845258
Patent No. 6183976
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Lodes, Michael J.
APPLICANT: Houghton, Raymond
APPLICANT: Sleath, Paul R.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS
TITLE OF INVENTION: AND TREATMENT OF B. MICROTI INFECTION
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/845,258
FILING DATE: 24-APR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.426C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 431 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-845-258-34

Query Match 33.3%; Score 5; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STEPS 12
Db 70 STEPS 74

RESULT 116
US-08-990-571-34
Sequence 34, Application US/08990571
Patent No. 6214971
GENERAL INFORMATION:
APPLICANT: Reed, Steven G. et al.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND TREATMENT OF B
NUMBER OF SEQUENCES: 79
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/990,571
FILING DATE: 11-DEC-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.426C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 431 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-990-571-34

Query Match 33.3%; Score 5; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STEPS 12
Db 70 STEPS 74

RESULT 117
US-08-723-142A-34
Sequence 34, Application US/08723142A
Patent No. 6306396
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Lodes, Michael J.
APPLICANT: Houghton, Raymond
APPLICANT: Sleath, Paul R.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS
TITLE OF INVENTION: AND TREATMENT OF B. MICROTI INFECTION
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY
STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/723,142A
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: MAKI, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.426
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 431 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-723-142A-34

Query Match 33.3%; Score 5; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPES 12
|||||
DB 70 STPES 74

RESULT 118
US-09-528-784A-34
Sequence 34, Application US/09528784A
Patent No. 6451315
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Lodes, Michael J.
APPLICANT: Houghton, Raymond L.
APPLICANT: Sleath, Paul R.
APPLICANT: McNeill, Patricia D.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS
FILE REFERENCE: 210121.426C4
CURRENT APPLICATION NUMBER: US/09/528,784A
CURRENT FILING DATE: 2000-03-17
NUMBER OF SEQ ID NOS: 90
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 34
LENGTH: 431
TYPE: PRT
ORGANISM: Babesia microti
US-09-528-784A-34

Query Match 33.3%; Score 5; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPES 12
|||||
DB 70 STPES 74

RESULT 119
US-08-261-110A-2
Sequence 2, Application US/08261110A
Patent No. 5674992

GENERAL INFORMATION:
APPLICANT: JACENDORF, ANDRE
APPLICANT: CERUTTI, HERIBERTO
TITLE OF INVENTION: CDNA ENCODING A RECA HOMOLOG IN
TITLE OF INVENTION: EUKARYOTES
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP
STREET: CLINTON SQUARE, P.O. BOX 1051
CITY: ROCHESTER
STATE: NEW YORK
COUNTRY: USA
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/261,110A
FILING DATE: 16-JUN-1994
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/938,332
FILING DATE: 28-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: TIMIAN, SUSAN J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 19603/10231(D-1292A)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1636
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 438 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-261-110A-2

Query Match 33.3%; Score 5; DB 1; Length 438;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 HRSTP 10
|||||
DB 410 HRSTP 414

RESULT 120
US-07-872-678A-47
Sequence 47, Application US/07872678A
Patent No. 5541060
GENERAL INFORMATION:
APPLICANT: Bell, Graeme, et al.
TITLE OF INVENTION: DETECTION OF EARLY-ONSET
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: Post Office Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/872,678A

FILING DATE: 22-APRIL-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coughlin, Daniel F.
REGISTRATION NUMBER: 36,111
REFERENCE/DOCKET NUMBER: ARCD016
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713-787-1400
TELEFAX: 713-789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 465 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-872-678A-47

Query Match 33.3%; Score 5; DB 1; Length 465;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
|||||
DB 63 RSTPE 67

RESULT 121
US-08-588-983-20
Sequence 20, Application US/08588983
Patent No. 5854067
GENERAL INFORMATION:
APPLICANT: Christopher B. Newgard, et al.
TITLE OF INVENTION: Methods and Compositions
TITLE OF INVENTION: for Inhibiting Hexokinase
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P. O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/588,983
FILING DATE: Concurrently herewith
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Fussey, Shelley P.M.
REGISTRATION NUMBER: 39,458
REFERENCE/DOCKET NUMBER: UTSD:424/FUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
TELEX: n/a
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 465 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-588-983-20

Query Match 33.3%; Score 5; DB 2; Length 465;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
|||||
DB 63 RSTPE 67

RESULT 122
US-08-588-983-22
Sequence 22, Application US/08588983
Patent No. 5854067
GENERAL INFORMATION:
APPLICANT: Christopher B. Newgard, et al.
TITLE OF INVENTION: Methods and Compositions
TITLE OF INVENTION: for Inhibiting Hexokinase
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P. O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/588,983
FILING DATE: Concurrently herewith
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Fussey, Shelley P.M.
REGISTRATION NUMBER: 39,458
REFERENCE/DOCKET NUMBER: UTSD:424/FUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
TELEX: n/a
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 465 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-588-983-22

Query Match 33.3%; Score 5; DB 2; Length 465;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
|||||
DB 63 RSTPE 67

RESULT 123
US-08-588-976-20
Sequence 20, Application US/08588976
Patent No. 5891717
GENERAL INFORMATION:
APPLICANT: Christopher B. Newgard, et al.
TITLE OF INVENTION: Methods and Compositions for
TITLE OF INVENTION: Inhibiting Hexokinase
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P. O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/588,976
FILING DATE: Concurrently herewith
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fussey, Shelley P.M.
REGISTRATION NUMBER: 39,458
REFERENCE/DOCKET NUMBER: UTSD:481/FUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
TELEX: n/a
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 465 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-588-976-20

Query Match 33.3%; Score 5; DB 2; Length 465;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
|||||
Db 63 RSTPE 67

RESULT 124
US-08-588-976-22
Sequence 22, Application US/08588976
Patent No. 5891717
GENERAL INFORMATION:
APPLICANT: Christopher B. Newgard, et al.
TITLE OF INVENTION: Methods and Compositions for
TITLE OF INVENTION: Inhibiting Hexokinase
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/588,976
FILING DATE: Concurrently herewith
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fussey, Shelley P.M.
REGISTRATION NUMBER: 39,458
REFERENCE/DOCKET NUMBER: UTSD:481/FUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
TELEX: n/a
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 465 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-588-976-22

Query Match 33.3%; Score 5; DB 2; Length 465;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
|||||
Db 63 RSTPE 67

RESULT 125
US-09-347-878-44
Sequence 44, Application US/09347878C
Patent No. 6376210
GENERAL INFORMATION:
APPLICANT: Yuan, Chong
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES
FILE REFERENCE: 25885-1651
CURRENT APPLICATION NUMBER: US/09/347,878C
CURRENT FILING DATE: 1999-07-06
NUMBER OF SEQ ID NOS: 75
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 44
LENGTH: 465
TYPE: PRT
ORGANISM: Homo sapiens
US-09-347-878-44

Query Match 33.3%; Score 5; DB 4; Length 465;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
|||||
Db 63 RSTPE 67

RESULT 126
US-08-249-112-4
Sequence 4, Application US/08249112
Patent No. 5527703
GENERAL INFORMATION:
APPLICANT: Cully, Doris F.
APPLICANT: Arena, Joseph P.
APPLICANT: Liu, Ken K.
APPLICANT: Vassiliadis, Demetrios
TITLE OF INVENTION: DNA ENCODING GLUTAMATE GATED CHLORIDE
TITLE OF INVENTION: CHANNELS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wallen, John W.
STREET: 126 E. Lincoln Ave., P.O. Box 2000
CITY: Rahway
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07065
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/249,112
FILING DATE: 25-MAY-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Wallen, John W.
REGISTRATION NUMBER: 35,403
REFERENCE/DOCKET NUMBER: 19194
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-3905
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:

LENGTH: 487 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-249-112-4

Query Match
 Best Local Similarity 33.3%; Score 5; DB 1; Length 487;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STEPS 12
 DB 432 STEPS 436

RESULT 127
 PCT-US95-06556-4
 ; Sequence 4, Application PC/TUS9506556
 ; GENERAL INFORMATION:
 ; APPLICANT: Cully, Doris F.
 ; APPLICANT: Arena, Joseph P.
 ; APPLICANT: Liu, Ken K.
 ; APPLICANT: Vassiliadis, Demetrios
 ; TITLE OF INVENTION: DNA ENCODING GLUTAMATE GATED CHLORIDE
 ; TITLE OF INVENTION: CHANNELS
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Wallen, John W.
 ; STREET: 126 E. Lincoln Ave., P.O. Box 2000
 ; CITY: Rahway
 ; STATE: New Jersey
 ; COUNTRY: U.S.A.
 ; ZIP: 07065
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US95/06556
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/249,112
 ; FILING DATE: 25-MAY-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Wallen, John W.
 ; REGISTRATION NUMBER: 35,403
 ; REFERENCE/DOCKET NUMBER: 19194
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (908) 594-3905
 ; TELEFAX: (908) 594-4720
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 487 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; PCT-US95-06556-4

Query Match
 Best Local Similarity 33.3%; Score 5; DB 5; Length 487;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STEPS 12
 DB 432 STEPS 436

RESULT 128
 US-09-426-568A-2

; Sequence 2, Application US/09426568A
 ; Patent No. 6348643
 ; GENERAL INFORMATION:
 ; APPLICANT: Kakefuda, Genichi
 ; APPLICANT: Costello, Colleen
 ; APPLICANT: Sun, Ming
 ; APPLICANT: Hu, Weiming
 ; TITLE OF INVENTION: Genes and Vectors for Confering Herbicide Resistance
 ; TITLE OF INVENTION: In Plants
 ; FILE REFERENCE: 008103/195497
 ; CURRENT APPLICATION NUMBER: US/09/426,568A
 ; CURRENT FILING DATE: 1999-10-22
 ; PRIOR APPLICATION NUMBER: 60/106,239
 ; PRIOR FILING DATE: 1998-10-29
 ; NUMBER OF SEQ ID NOS: 11
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 491
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis sp.
 US-09-426-568A-2

Query Match
 Best Local Similarity 33.3%; Score 5; DB 4; Length 491;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRA 14
 DB 177 PESRA 181

RESULT 129
 US-09-431-470-4
 ; Sequence 4, Application US/09431470
 ; Patent No. 6435249
 ; GENERAL INFORMATION:
 ; APPLICANT: Duvick, Jon
 ; APPLICANT: Simmons, Carl R.
 ; APPLICANT: Crasta, Oswald R.
 ; APPLICANT: Folkerts, Otto
 ; TITLE OF INVENTION: The Use of Beta-Glucosidase to Enhance
 ; TITLE OF INVENTION: Disease Resistance to Insects in Crop Plants.
 ; FILE REFERENCE: 5718-43
 ; CURRENT APPLICATION NUMBER: US/09/431,470
 ; CURRENT FILING DATE: 1999-11-01
 ; PRIOR APPLICATION NUMBER: 60/107,920
 ; PRIOR FILING DATE: 1998-11-10
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 563
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 US-09-431-470-4

Query Match
 Best Local Similarity 33.3%; Score 5; DB 4; Length 563;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGP 5
 DB 19 SHLGP 23

RESULT 130
 US-08-399-646-2
 ; Sequence 2, Application US/08399646
 ; Patent No. 5556781
 ; GENERAL INFORMATION:
 ; APPLICANT: KUBOTA, Michio
 ; APPLICANT: TSUSAKI, Keiji
 ; APPLICANT: HATTORI, Kazuko
 ; APPLICANT: SUGIMOTO, Toshiyuki

;; TITLE OF INVENTION: DNA ENCODING ENZYME, RECOMBINANT DNA AND
;; TITLE OF INVENTION: ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
;; NUMBER OF SEQUENCES: 18
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: BROWDY AND NEIMARK
;; STREET: 419 Seventh Street, N.W., Suite 300
;; CITY: Washington
;; STATE: D.C.
;; COUNTRY: USA
;; ZIP: 20004
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/399,646
;; FILING DATE: 07-MAR-1995
;;
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 59834
;; FILING DATE: 07-MAR-1994
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 59840
;; FILING DATE: 07-MAR-1994
;;
;; ATTORNEY/AGENT INFORMATION:
;; NAME: BROWDY, Roger L.
;; REGISTRATION NUMBER: 25,618
;; REFERENCE/DOCKET NUMBER: KUBOTA=5
;;
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-628-5197
;; TELEFAX: 202-737-3528
;; TELEX: 248633
;;
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 589 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;;
;; US-08-399-646-2
;;
Query Match 33.3%; Score 5; DB 1; Length 589;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 RSTPE 11
DB 511 RSTPE 515
;;
RESULT 131
US-08-607-321-2
;; Sequence 2, Application US/08607321
;; Patent No. 5716813
;;
;; GENERAL INFORMATION:
;; APPLICANT: KUBOTA, Michio
;; APPLICANT: TSUSAKI, Keiji
;; APPLICANT: HATTORI, Kazuko
;; APPLICANT: SUGIMOTO, Toshiyuki
;; TITLE OF INVENTION: DNA ENCODING ENZYME, RECOMBINANT DNA AND
;; TITLE OF INVENTION: ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
;; NUMBER OF SEQUENCES: 18
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: BROWDY AND NEIMARK
;; STREET: 419 Seventh Street, N.W., Suite 300
;; CITY: Washington
;; STATE: D.C.
;; COUNTRY: USA
;; ZIP: 20004
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: Patentin Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/607,321
;; FILING DATE: 26-FEB-1996
;;
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/399,646
;; FILING DATE: 07-MAR-1995
;; APPLICATION NUMBER: JP 59834
;; FILING DATE: 07-MAR-1994
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 59840
;; FILING DATE: 07-MAR-1994
;;
;; ATTORNEY/AGENT INFORMATION:
;; NAME: BROWDY, Roger L.
;; REGISTRATION NUMBER: 25,618
;; REFERENCE/DOCKET NUMBER: KUBOTA=5
;;
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-628-5197
;; TELEFAX: 202-737-3528
;; TELEX: 248633
;;
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 589 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;;
;; US-08-607-321-2
;;
Query Match 33.3%; Score 5; DB 1; Length 589;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 RSTPE 11
DB 511 RSTPE 515
;;
RESULT 132
US-08-961-240-2
;; Sequence 2, Application US/08961240
;; Patent No. 5830715
;;
;; GENERAL INFORMATION:
;; APPLICANT: KUBOTA, Michio
;; APPLICANT: TSUSAKI, Keiji
;; APPLICANT: HATTORI, Kazuko
;; APPLICANT: SUGIMOTO, Toshiyuki
;; TITLE OF INVENTION: DNA ENCODING ENZYME, RECOMBINANT DNA AND
;; TITLE OF INVENTION: ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
;; NUMBER OF SEQUENCES: 18
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: BROWDY AND NEIMARK
;; STREET: 419 Seventh Street, N.W., Suite 300
;; CITY: Washington
;; STATE: D.C.
;; COUNTRY: USA
;; ZIP: 20004
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/961,240
;; FILING DATE: 30-OCT-1997
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/399,646
;; FILING DATE: 07-MAR-1995
;; APPLICATION NUMBER: JP 59834
;; FILING DATE: 07-MAR-1994
;; PRIOR APPLICATION DATA: JP 59840

FILING DATE: 07-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: KUBOTA-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ. ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 589 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-961-240-2

Query Match 33.3%; Score 5; DB 2; Length 589;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPE 11
DB 511 RSTPE 515

RESULT 133
US-08-605-501-2
Sequence 2, Application US/08605501
Patent No. 5834287
GENERAL INFORMATION:
APPLICANT: KUBOTA, Michio
APPLICANT: TSUSAKI, Keiji
APPLICANT: HATTORI, Kazuko
APPLICANT: SUGIMOTO, Toshiyuki
TITLE OF INVENTION: DNA ENCODING ENZYME, RECOMBINANT DNA AND
TITLE OF INVENTION: ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NETMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/605,501
FILING DATE: 26-FEB-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/399,646
FILING DATE: 07-MAR-1995
APPLICATION NUMBER: JP 59834
FILING DATE: 07-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 59840
FILING DATE: 07-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: KUBOTA-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ. ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 589 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-605-501-2

Query Match 33.3%; Score 5; DB 2; Length 589;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPE 11
DB 511 RSTPE 515

RESULT 134
US-08-399-646-12
Sequence 12, Application US/08399646
Patent No. 5556781
GENERAL INFORMATION:
APPLICANT: KUBOTA, Michio
APPLICANT: TSUSAKI, Keiji
APPLICANT: HATTORI, Kazuko
APPLICANT: SUGIMOTO, Toshiyuki
TITLE OF INVENTION: DNA ENCODING ENZYME, RECOMBINANT DNA AND
TITLE OF INVENTION: ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NETMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/399,646
FILING DATE: 07-MAR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 59834
FILING DATE: 07-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 59840
FILING DATE: 07-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: KUBOTA-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ. ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 596 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-399-646-12

Query Match 33.3%; Score 5; DB 1; Length 596;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPE 11
DB 518 RSTPE 522

RESULT 135
US-08-607-321-12
; Sequence 12, Application US/08607321
; Patent No. 5716813
; GENERAL INFORMATION:
; APPLICANT: KUBOTA, Michio
; APPLICANT: TSUSAKI, Keiji
; APPLICANT: HATTORI, Kazuko
; APPLICANT: SUGIMOTO, Toshiyuki
; TITLE OF INVENTION: DNA ENCODING ENZYME, RECOMBINANT DNA AND
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/607,321
; FILING DATE: 26-FEB-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/399,646
; FILING DATE: 07-MAR-1995
; APPLICATION NUMBER: JP 59834
; FILING DATE: 07-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 59840
; FILING DATE: 07-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: KUBOTA=5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 596 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-607-321-12

Query Match 33.3%; Score 5; DB 1; Length 596;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
|||||
DB 518 RSTPE 522

RESULT 136
US-08-961-240-12
; Sequence 12, Application US/08961240
; Patent No. 5830715
; GENERAL INFORMATION:
; APPLICANT: KUBOTA, Michio
; APPLICANT: TSUSAKI, Keiji
; APPLICANT: HATTORI, Kazuko
; APPLICANT: SUGIMOTO, Toshiyuki
; TITLE OF INVENTION: DNA ENCODING ENZYME, RECOMBINANT DNA AND
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30

CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,240
FILING DATE: 30-OCT-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/399,646
FILING DATE: 07-MAR-1995
APPLICATION NUMBER: JP 59834
FILING DATE: 07-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 59840
FILING DATE: 07-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: KUBOTA=5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 596 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-961-240-12

Query Match 33.3%; Score 5; DB 2; Length 596;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
|||||
DB 518 RSTPE 522

RESULT 137
US-08-605-501-12
; Sequence 12, Application US/08605501
; Patent No. 5834287
; GENERAL INFORMATION:
; APPLICANT: KUBOTA, Michio
; APPLICANT: TSUSAKI, Keiji
; APPLICANT: HATTORI, Kazuko
; APPLICANT: SUGIMOTO, Toshiyuki
; TITLE OF INVENTION: DNA ENCODING ENZYME, RECOMBINANT DNA AND
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/605,501
FILING DATE: 26-FEB-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/399,646
FILING DATE: 07-MAR-1995
APPLICATION NUMBER: JP 59834
FILING DATE: 07-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 59840
FILING DATE: 07-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: KUBOTA-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ. ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 596 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-605-501-12

Query Match 33.3%; Score 5; DB 2; Length 596;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
|||||
Db 518 RSTPE 522

RESULT 138
PCT-US93-00031-19
Sequence 19, Application PC/TUS9300031
GENERAL INFORMATION:
APPLICANT: Osborn, Laurelee
REGISTRATION/DOCKET NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 92,306-A; D001 CIP PCT
TITLE OF INVENTION: ANTIBODIES RECOGNIZING THE FOURTH
TITLE OF INVENTION: IMMUNOGLOBULIN-LIKE DOMAIN OF VCAM1
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Allegretti & Witcoff, Ltd.
STREET: 10 South Wacker Drive
CITY: Chicago
STATE: IL
COUNTRY: US
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/00031
FILING DATE: 19930112
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: McNicholas, Janet M.
REGISTRATION/DOCKET NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 92,306-A; D001 CIP PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 715-1234
TELEFAX: (312) 715-1234
INFORMATION FOR SEQ. ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 643 amino acids
TYPE: AMINO ACID

TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US93-00031-19

Query Match 33.3%; Score 5; DB 5; Length 643;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESR 13
|||||
Db 30 TPESR 34

RESULT 139
PCT-US93-00031-21
Sequence 21, Application PC/TUS9300031
GENERAL INFORMATION:
APPLICANT: Osborn, Laurelee
REGISTRATION/DOCKET NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 92,306-A; D001 CIP PCT
TITLE OF INVENTION: ANTIBODIES RECOGNIZING THE FOURTH
TITLE OF INVENTION: IMMUNOGLOBULIN-LIKE DOMAIN OF VCAM1
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Allegretti & Witcoff, Ltd.
STREET: 10 South Wacker Drive
CITY: Chicago
STATE: IL
COUNTRY: US
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/00031
FILING DATE: 19930112
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: McNicholas, Janet M.
REGISTRATION/DOCKET NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 92,306-A; D001 CIP PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 715-1234
TELEFAX: (312) 715-1234
INFORMATION FOR SEQ. ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 644 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US93-00031-21

Query Match 33.3%; Score 5; DB 5; Length 644;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESR 13
|||||
Db 30 TPESR 34

RESULT 140
US-09-009-490A-91
Sequence 91, Application US/09009490A
Patent No. 6300491
GENERAL INFORMATION:
APPLICANT: Bennett and Mirabelli
TITLE OF INVENTION: Oligonucleotide Modulation
TITLE OF INVENTION: Of Cell Adhesion
NUMBER OF SEQUENCES: 95
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Law Office of Jane Massey Licata

STREET: 66 East Main Street
CITY: Marlton
STATE: NJ
COUNTRY: USA
ZIP: 08053
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
COMPUTER: IBM PS/2
OPERATING SYSTEM: Windows 95
SOFTWARE: WORDPERECT 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/009,490A
FILING DATE: January 20, 1998
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 440,740
FILING DATE: May 12, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 063,167
FILING DATE: May 17, 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 969,151
FILING DATE: February 10, 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 007,997
FILING DATE: January 20, 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 939,855
FILING DATE: September 2, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 567,286
FILING DATE: August 14, 1990
ATTORNEY/AGENT INFORMATION:
NAME: Jane Massey Licata
REGISTRATION NUMBER: 32,257
REFERENCE/DOCKET NUMBER: ISPH-0268
TELECOMMUNICATION INFORMATION:
TELEPHONE: (609) 810-1515
TELEFAX: (609) 810-1454
INFORMATION FOR SEQ ID NO: 91:
SEQUENCE CHARACTERISTICS:
LENGTH: 647
TYPE: Amino Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
ANTI-SENSE: NO
US-09-009-490A-91

Query Match 33.3%; Score 5; DB 4; Length 647;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 TPESR 13
Db 30 TPESR 34

RESULT 141
US-08-482-073-5
Sequence 5, Application US/08482073
GENERAL INFORMATION:
APPLICANT: Hession, Catherine A.
APPLICANT: Lobb, Roy R.
APPLICANT: Goeltz, Susan E.
APPLICANT: Osborn, Laurelee
APPLICANT: Benjamin, Christopher D.
APPLICANT: Rosa, Margaret D.
TITLE OF INVENTION: ENOTOHELIAL CELL-LEUKOCYTE ADHESION
TITLE OF INVENTION: MOLECULES (ELAMS) AND MOLECULES INVOLVED IN LEUKOCYTE
TITLE OF INVENTION: ADHESION (MILAS)
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Neave
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10020
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,073
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/486,336
FILING DATE:
APPLICATION NUMBER: US 07/608298
FILING DATE: 31-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US 90/02357
FILING DATE: 27-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/452675
FILING DATE: 18-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/359516
FILING DATE: 01-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/354151
FILING DATE: 28-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Haley Jr., James F.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: B124C1P4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 596-9000
TELEFAX: (212) 596-9090
TELEX: 14-8367
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 647 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-482-073-5

Query Match 33.3%; Score 5; DB 4; Length 647;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 TPESR 13
Db 30 TPESR 34

RESULT 142
PCT-US93-00031-11
Sequence 11, Application PC/TUS9300031
GENERAL INFORMATION:
APPLICANT: Osborn, Laurelee
APPLICANT: Benjamin, Christopher D.
TITLE OF INVENTION: ANTIBODIES RECOGNIZING THE FOURTH
TITLE OF INVENTION: IMMUNOGLOBULIN-LIKE DOMAIN OF VCAM1
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Allegretti & Witcoff, Ltd.
STREET: 10 South Wacker Drive
CITY: Chicago
STATE: IL
COUNTRY: US
ZIP: 60606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/00031
FILING DATE: 19930112
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: McNicholas, Janet M.
REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 92,306-A; D001 CIP PCT
TELEPHONE: (312) 715-1000
TELEFAX: (312) 715-1234
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 647 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US93-00031-11

Query Match 33.3%; Score 5; DB 5; Length 647;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESR 13
Db 30 TPESR 34

RESULT 143
US-08-261-304-7
Sequence 7, Application US/08261304
Patent No. 5708147
GENERAL INFORMATION:
APPLICANT: Cybulsky, Myron I.
APPLICANT: Gimbrone, Michael A.
TITLE OF INVENTION: Mononuclear Leukocyte Directed
TITLE OF INVENTION: Endothelial Adhesion Molecule Associated with
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1225 Connecticut Avenue, N.W.
CITY: Washington
STATE: District of Columbia
COUNTRY: United States of America
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Ascii
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/261,304
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/649,565
FILING DATE: 01-FEB-1991
APPLICATION NUMBER: U.S. 07/487,038
FILING DATE: 02-MAR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Potter, Jane E. R.
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 0627.2100004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 833-7533

INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 662 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-261-304-7

Query Match 33.3%; Score 5; DB 1; Length 662;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESR 13
Db 6 TPESR 10

RESULT 144
US-08-205-018-2
Sequence 2, Application US/08205018
Patent No. 5554523
GENERAL INFORMATION:
APPLICANT: Reddy, Usharani R.
TITLE OF INVENTION: No. 5554523el Protein Kinase, Nucleic Acid
TITLE OF INVENTION: Sequences Encoding the Same and Methods Related
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and
ADDRESSEE: No. 5554523rls
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/205,018
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Gaumont, Rebecca R.
REGISTRATION NUMBER: 35,152
REFERENCE/DOCKET NUMBER: CH-0488
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 668 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-205-018-2

Query Match 33.3%; Score 5; DB 1; Length 668;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRA 14
Db 65 PESRA 69

RESULT 145
US-08-424-424B-2
Sequence 2, Application US/08424424B
Patent No. 5759854

GENERAL INFORMATION:
APPLICANT: LI, ET AL.
TITLE OF INVENTION: Neurotransmitter Transporter
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/424,424B
FILING DATE: APRIL 21, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/05363
FILING DATE: MAY 25, 1996
ATTORNEY/AGENT INFORMATION:
NAME: MULLINS, J.G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-308
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 727 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-08-424-424B-2

Query Match 33.3%; Score 5; DB 1; Length 727;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPES 12
|||||
DB 721 STPES 725

RESULT 146
PCT-US94-05363A-2
Sequence 2, Application PC/TUS9405363A
GENERAL INFORMATION:
APPLICANT: LI, ET AL.
TITLE OF INVENTION: Neurotransmitter Transporter
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/05363A
FILING DATE: SUBMITTED HEREMITH
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: FERRARO, GREGORY D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-118
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 727 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
PCT-US94-05363A-2

Query Match 33.3%; Score 5; DB 5; Length 727;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPES 12
|||||
DB 721 STPES 725

RESULT 147
US-08-696-944-20
Sequence 20, Application US/08696944
Patent No. 5981831
GENERAL INFORMATION:
APPLICANT: SUMANT CHENGAPPA
APPLICANT: SUSAN A. HELLYER
APPLICANT: JOHN S. REID
TITLE OF INVENTION: NO. 5981831el Exo-(1-4)-Beta-D Galactanase
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/696,944
FILING DATE: 23-AUG-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB95/00372
FILING DATE: 23-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9403423.8
FILING DATE: 23-FEB-1994
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 731 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-696-944-20

Query Match 33.3%; Score 5; DB 2; Length 731;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 STPES 11

DB 56 RSTPE 60

|||||

RESULT 148

US-08-706-216-2

; Sequence 2, Application US/08706216

; Patent No. 6140098

GENERAL INFORMATION:

APPLICANT: Balasubramanian, Sritam

APPLICANT: Ford, John

APPLICANT: Gorman, Daniel M.

APPLICANT: Zurawski, Gerard

TITLE OF INVENTION: MAMMALIAN PROTEASES; RELATED REAGENTS

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: DNAX Research Institute

STREET: 901 California Avenue

CITY: Palo Alto

STATE: California

COUNTRY: USA

ZIP: 94304-1104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/706,216

FILING DATE: 30-AUG-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Ching, Edwin P.

REGISTRATION NUMBER: 34,090

REFERENCE/DOCKET NUMBER: DX0613

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-852-9196

TELEFAX: 415-496-1200

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 734 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-706-216-2

Query Match 33.3%; Score 5; DB 4; Length 734;

Best Local Similarity 100.0%; Pred. No. 6.6e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPHR 7

DB 140 LGPHR 144

RESULT 149

US-09-641-741-2

; Sequence 2, Application US/09641741

; Patent No. 642015

GENERAL INFORMATION:

APPLICANT: Kerry E. Quinn

APPLICANT: Curagen Corporation

TITLE OF INVENTION: Aortic Carboxypeptidase-Like Proteins and Nucleic Acids

TITLE OF INVENTION: encoding Same

FILE REFERENCE: 15966-581

CURRENT APPLICATION NUMBER: US/09/641,741

CURRENT FILING DATE: 2000-08-18

PRIOR APPLICATION NUMBER: 60/159,613

PRIOR FILING DATE: 1999-10-14

PRIOR APPLICATION NUMBER: 60/175,534

PRIOR FILING DATE: 2000-01-12

PRIOR APPLICATION NUMBER: 60/224,086

PRIOR FILING DATE: 2000-08-09

; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO: 2

; LENGTH: 734

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-641-741-2

Query Match 33.3%; Score 5; DB 4; Length 734;

Best Local Similarity 100.0%; Pred. No. 6.6e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPHR 7

DB 140 LGPHR 144

RESULT 150

PCT-US93-00031-13

; Sequence 13, Application PC/TUS9300031

GENERAL INFORMATION:

APPLICANT: Osborn, Laurelee

APPLICANT: Benjamin, Christopher D.

TITLE OF INVENTION: ANTIBODIES RECOGNIZING THE FOURTH

TITLE OF INVENTION: IMMUNOGLOBULIN-LIKE DOMAIN OF VCAM1

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSEE: Allegretti & Witcoff, Ltd.

STREET: 10 South Wacker Drive

CITY: Chicago

STATE: IL

COUNTRY: US

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/00031

FILING DATE: 19930112

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: McNicholas, Janet M.

REGISTRATION NUMBER: 32,918

REFERENCE/DOCKET NUMBER: 92,306-A; D001 CIP PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: (312) 715-1000

TELEFAX: (312) 715-1234

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 735 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US93-00031-13

Query Match 33.3%; Score 5; DB 5; Length 735;

Best Local Similarity 100.0%; Pred. No. 6.6e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 TPESR 13

DB 30 TPESR 34

RESULT 151

PCT-US93-00031-15

; Sequence 15, Application PC/TUS9300031

GENERAL INFORMATION:

APPLICANT: Osborn, Laurelee

APPLICANT: Benjamin, Christopher D.

TITLE OF INVENTION: ANTIBODIES RECOGNIZING THE FOURTH

TITLE OF INVENTION: IMMUNOGLOBULIN-LIKE DOMAIN OF VCAM1
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Allegretti & Wilcoff, Ltd.
STREET: 10 South Wacker Drive
CITY: Chicago
STATE: IL
COUNTRY: US
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/00031
FILING DATE: 19930112
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: McNicholas, Janet M.
REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 92,306-A; D001 CIP PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 715-1000
TELEFAX: (312) 715-1234
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 736 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US93-00031-15

Query Match 33.3%; Score 5; DB 5; Length 736;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 9 TPESR 13
Db 30 TPESR 34

RESULT 152
US-08-482-073-6
Sequence 6, Application US/08482073
Patent No. 6307025
GENERAL INFORMATION:
APPLICANT: Hession, Catherine A.
APPLICANT: Lobb, Roy R.
APPLICANT: Goeltz, Susan E.
APPLICANT: Osborn, Laurelee
APPLICANT: Benjamin, Christopher D.
APPLICANT: Rosa, Margaret D.
TITLE OF INVENTION: ENDOTHELIAL CELL-LEUKOCYTE ADHESION
TITLE OF INVENTION: MOLECULES (ELAMS) AND MOLECULES INVOLVED IN LEUKOCYTE
TITLE OF INVENTION: ADHESION (MILAS)
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Neave
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10020
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,073
FILING DATE:
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/486,336
FILING DATE:
APPLICATION NUMBER: US 07/608298
FILING DATE: 31-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US 90/02357
FILING DATE: 27-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/452675
FILING DATE: 18-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/359516
FILING DATE: 01-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/354151
FILING DATE: 28-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Haley Jr., James F.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: B124CIP4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 596-9000
TELEFAX: (212) 596-9090
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 739 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-482-073-6

Query Match 33.3%; Score 5; DB 4; Length 739;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 9 TPESR 13
Db 30 TPESR 34

RESULT 153
PCT-US93-00031-9
Sequence 9, Application PC/TUS9300031
GENERAL INFORMATION:
APPLICANT: Osborn, Laurelee
APPLICANT: Benjamin, Christopher D.
TITLE OF INVENTION: ANTIBODIES RECOGNIZING THE FOURTH
TITLE OF INVENTION: IMMUNOGLOBULIN-LIKE DOMAIN OF VCAM1
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Allegretti & Wilcoff, Ltd.
STREET: 10 South Wacker Drive
CITY: Chicago
STATE: IL
COUNTRY: US
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/00031
FILING DATE: 19930112
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: McNicholas, Janet M.
REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 92,306-A; D001 CIP PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 715-1000

TELEFAX: (312) 715-1234
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 739 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US93-00031-9

Query Match 33.3%; Score 5; DB 5; Length 739;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESR 13
DB 30 TPESR 34

RESULT 154
PCT-US93-00031-17

; Sequence 17, Application PC/TUS9300031
; GENERAL INFORMATION:
; APPLICANT: Osborn, Laurelee
; TITLE OF INVENTION: ANTIBODIES RECOGNIZING THE FOURTH
; TITLE OF INVENTION: IMMUNOGLOBULIN-LIKE DOMAIN OF VCAM1
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive
; CITY: Chicago
; STATE: IL
; COUNTRY: US
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/00031
; FILING DATE: 19930112
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: McNicholas, Janet M.
; REGISTRATION NUMBER: 32,918
; REFERENCE/DOCKET NUMBER: 92,306-A; D001 CIP PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 715-1000
; TELEFAX: (312) 715-1234
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 740 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US93-00031-17

Query Match 33.3%; Score 5; DB 5; Length 740;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESR 13
DB 30 TPESR 34

RESULT 155

US-09-342-648-2
; Sequence 2, Application US/09342648
; Patent No. 6248584
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.

; APPLICANT: Odell, Joan
; APPLICANT: Rafalski, Antoni
; TITLE OF INVENTION: Transcription Coactivators
; FILE REFERENCE: BB-1169-B
; CURRENT APPLICATION NUMBER: US/09/342,648
; CURRENT FILING DATE: 1999-06-29
; EARLIER APPLICATION NUMBER: 60/092,659
; EARLIER FILING DATE: July 13, 1998
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 755
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (179)
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (185)
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (241)
US-09-342-648-2

Query Match 33.3%; Score 5; DB 4; Length 755;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPEs 12
DB 492 STPEs 496

RESULT 156

US-08-696-944-19
; Sequence 19, Application US/08696944
; Patent No. 5981831
; GENERAL INFORMATION:
; APPLICANT: Sumant CHENGAPPA
; APPLICANT: Susan A. HELLYER
; APPLICANT: John S. REID
; TITLE OF INVENTION: NO. 5981831el EXO-(1-4)-Beta-D Galactanase
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/696,944
; FILING DATE: 23-AUG-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/00372
; FILING DATE: 23-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9403423.8
; FILING DATE: 23-FEB-1994
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 838 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-696-944-19

Query Match 33.3%; Score 5; DB 2; Length 838;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
|||||
DB 52 RSTPE 56

RESULT 157

US-08-395-580-2
Sequence 2, Application US/08395580
Patent No. 5676945
GENERAL INFORMATION:
APPLICANT: Usharani R. Reddy, David Pleasure and the Children's
APPLICANT: Hospital of Philadelphia
TITLE OF INVENTION: No. 5676945e1 Protein Kinase, Nucleic Acid
NUMBER OF SEQUENCES: 4
SEQUENCES Encoding the Same and Methods Related Thereto
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and No. 5676945r1s
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: U.S.A.
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch disk, 720 Kb
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/395,580
FILING DATE: herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/205,018
FILING DATE: 01-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Rebecca L. Ralph (formerly Gaumond)
REGISTRATION NUMBER: 35,152
REFERENCE/DOCKET NUMBER: CH-0488
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 859 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-395-580-2

Query Match 33.3%; Score 5; DB 1; Length 859;
Best Local Similarity 100.0%; Pred. No. 7.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRA 14
|||||
DB 65 PESRA 69

RESULT 158

PCT-US95-02792-2
Sequence 2, Application PC/TUS9502792
GENERAL INFORMATION:
APPLICANT: Usharani R. Reddy, David Pleasure and the Children's
APPLICANT: Hospital of Philadelphia
TITLE OF INVENTION: Novel Protein Kinase, Nucleic Acid
NUMBER OF SEQUENCES: 4
SEQUENCES Encoding the Same and Methods Related Thereto
CORRESPONDENCE ADDRESS:

ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and Norris
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: U.S.A.
ZIP: 19103

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch disk, 720 Kb
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02792
FILING DATE: herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/205,018
FILING DATE: 01-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Rebecca L. Ralph (formerly Gaumond)
REGISTRATION NUMBER: 35,152
REFERENCE/DOCKET NUMBER: CH-0488
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 859 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-02792-2

Query Match 33.3%; Score 5; DB 5; Length 859;
Best Local Similarity 100.0%; Pred. No. 7.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRA 14
|||||
DB 65 PESRA 69

RESULT 159
US-09-150-460B-7
Sequence 7, Application US/09150460B
Patent No. 6190882
GENERAL INFORMATION:
APPLICANT: Lee, Cheng-Chi
APPLICANT: Aldrecht, Urs
APPLICANT: Elchele, Gregor
APPLICANT: Sun, Zhong Sheng
TITLE OF INVENTION: Mammalian Circadian Rhythm-Like Gene
FILE REFERENCE: D6039

CURRENT APPLICATION NUMBER: US/09/150,460B
CURRENT FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: US 60/058,256
PRIOR FILING DATE: 1997-09-09
NUMBER OF SEQ ID NOS: 21
SEQ ID NO 7
LENGTH: 875
TYPE: PRT
ORGANISM: artificial sequence
FEATURE:
OTHER INFORMATION: Peptide sequence of largest deduced open reading
OTHER INFORMATION: frame from RIGUI 6.6
US-09-150-460B-7

Query Match 33.3%; Score 5; DB 4; Length 875;
Best Local Similarity 100.0%; Pred. No. 7.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGP 5
|||||

Db 822 SHLGP 826

RESULT 160

US-08-574-959A-9
Sequence 9, Application US/08574959A
Patent No. 5962224

GENERAL INFORMATION:

APPLICANT: Jaekyoon Shin, Insil Joung, Ratna K. Vadlamudi
APPLICANT: and Jack L. Strominger
TITLE OF INVENTION: p62 POLYPEPTIDES, RELATED POLYPEPTIDES
TITLE OF INVENTION: AND USES THEREFOR
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109-1875

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/574,959A

FILING DATE: 19-DEC-95

ATTORNEY/AGENT INFORMATION:

NAME: Mandragouras, Amy E.

REGISTRATION NUMBER: 36,207

REFERENCE/DOCKET NUMBER: DFN-008

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-7400

TELEFAX: (617)227-5941

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 905 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-574-959A-9

Query Match 33.3%; Score 5; DB 2; Length 905;

Best Local Similarity 100.0%; Pred. No. 8e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 GPHRS 8

DB 429 GPHRS 433

RESULT 161

US-09-357-014-9

Sequence 9, Application US/09357014

Patent No. 6291645

GENERAL INFORMATION:

APPLICANT: Jaekyoon Shin, Insil Joung, Ratna K. Vadlamudi
APPLICANT: and Jack L. StromingerTITLE OF INVENTION: p62 POLYPEPTIDES, RELATED POLYPEPTIDES
TITLE OF INVENTION: AND USES THEREFOR

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD

STREET: 60 State Street, Suite 510

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109-1875

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/357,014

FILING DATE: 19-Jul-1999

PRIOR APPLICATION NUMBER: 08/574,959

APPLICATION DATA:

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Mandragouras, Amy E.

REGISTRATION NUMBER: 36,207

REFERENCE/DOCKET NUMBER: DFN-008

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-7400

TELEFAX: (617)227-5941

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 905 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 9:

US-09-357-014-9

Query Match 33.3%; Score 5; DB 4; Length 905;

Best Local Similarity 100.0%; Pred. No. 8e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 GPHRS 8

DB 429 GPHRS 433

RESULT 162

US-08-588-983-9

Sequence 9, Application US/08588983

Patent No. 5854067

GENERAL INFORMATION:

APPLICANT: Christopher B. Newgard, et al.

TITLE OF INVENTION: Methods and Compositions

TITLE OF INVENTION: for Inhibiting Hexokinase

NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee

STREET: P.O. Box 4433

CITY: Houston

STATE: TX

COUNTRY: US

ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/588,983

FILING DATE: Concurrently herewith

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Fussey, Shelley P.M.

REGISTRATION NUMBER: 39,458

REFERENCE/DOCKET NUMBER: US/08/588,983

TELECOMMUNICATION INFORMATION:

TELEPHONE: (512) 418-3000

TELEFAX: (512) 474-7577

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 919 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-588-983-9

Query Match 33.3%; Score 5; DB 2; Length 919;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPE 11
|||||
DB 517 RSTPE 521

RESULT 163
US-08-588-983-12

; Sequence 12, Application US/08588983
; Patent No. 5854067

; GENERAL INFORMATION:

; APPLICANT: Christopher B. Newgard, et al.

; TITLE OF INVENTION: Methods and Compositions
for Inhibiting Hexokinase

; NUMBER OF SEQUENCES: 43

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Arnold, White & Durkee

; STREET: P.O. Box 4433

; CITY: Houston

; STATE: TX

; COUNTRY: US

; ZIP: 77210

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/588,983

; FILING DATE: Concurrently herewith

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Fussey, Shelley P.M.

; REGISTRATION NUMBER: 39,458

; REFERENCE/DOCKET NUMBER: UTSD:424/FUS

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (512) 418-3000

; TELEFAX: (512) 474-7577

; TELEX: n/a

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 919 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; US-08-588-983-12

; Query Match 33.3%; Score 5; DB 2; Length 919;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPE 11
|||||

DB 517 RSTPE 521

RESULT 164
US-08-588-976-9

; Sequence 9, Application US/08588976
; Patent No. 5891717

; GENERAL INFORMATION:

; APPLICANT: Christopher B. Newgard, et al.

; TITLE OF INVENTION: Methods and Compositions for
Inhibiting Hexokinase

; NUMBER OF SEQUENCES: 43

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Arnold, White & Durkee

; STREET: P.O. Box 4433

; CITY: Houston

; STATE: TX

; COUNTRY: US
; ZIP: 77210

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/588,976

; FILING DATE: Concurrently herewith

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Fussey, Shelley P.M.

; REGISTRATION NUMBER: 39,458

; REFERENCE/DOCKET NUMBER: UTSD:481/FUS

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (512) 418-3000

; TELEFAX: (512) 474-7577

; TELEX: n/a

; INFORMATION FOR SEQ ID NO: 9:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 919 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; US-08-588-976-9

; Query Match 33.3%; Score 5; DB 2; Length 919;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPE 11
|||||

DB 517 RSTPE 521

RESULT 165
US-08-588-976-12

; Sequence 12, Application US/08588976
; Patent No. 5891717

; GENERAL INFORMATION:

; APPLICANT: Christopher B. Newgard, et al.

; TITLE OF INVENTION: Methods and Compositions for
Inhibiting Hexokinase

; NUMBER OF SEQUENCES: 43

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Arnold, White & Durkee

; STREET: P.O. Box 4433

; CITY: Houston

; STATE: TX

; COUNTRY: US

; ZIP: 77210

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/588,976

; FILING DATE: Concurrently herewith

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Fussey, Shelley P.M.

; REGISTRATION NUMBER: 39,458

; REFERENCE/DOCKET NUMBER: UTSD:481/FUS

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (512) 418-3000

; TELEFAX: (512) 474-7577

; TELEX: n/a

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 919 amino acids

; TYPE: amino acid

```
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-588-976-12

Query Match
Best Local Similarity 33.3%; Score 5; DB 2; Length 919;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPE 11
Db 517 RSTPE 521

RESULT 166
US-09-078-347A-1
; Sequence 1, Application US/09078347A
; Patent No. 6132968
; GENERAL INFORMATION:
; APPLICANT: Le, Xiao-Chun
; APPLICANT: Weinfield, Michael
; APPLICANT: Xing, James Z.
; TITLE OF INVENTION: Methods for Quantitating Low Level
; TITLE OF INVENTION: Modifications of Nucleotide Sequences
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/078,347A
; FILING DATE: 13-May-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: UALB-03283
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 940 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-09-078-347A-1

Query Match
Best Local Similarity 33.3%; Score 5; DB 4; Length 940;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRA 14
Db 719 PESRA 723

RESULT 167
US-09-651-656-101
; Sequence 101, Application US/09651656
; Patent No. 6340566
; GENERAL INFORMATION:
; APPLICANT: MCCUTHEN-MALONEY, SANDRA
; APPLICANT: LAWRENCE LIVERMORE NATIONAL LABORATORY
; TITLE OF INVENTION: DETECTION AND QUANTITATION OF SINGLE NUCLEOTIDE
```

```
; TITLE OF INVENTION: POLYMORPHISMS, DNA SEQUENCE VARIATIONS, DNA MUTATIONS,
; TITLE OF INVENTION: DNA DAMAGE AND DNA MISMATCHES
; FILE REFERENCE: IL-10689
; CURRENT APPLICATION NUMBER: US/09/651,656
; CURRENT FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/192,764
; PRIOR FILING DATE: 2000-03-28
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 101
; LENGTH: 940
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-651-656-101

Query Match
Best Local Similarity 33.3%; Score 5; DB 4; Length 940;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRA 14
Db 719 PESRA 723

RESULT 168
US-09-650-855-101
; Sequence 101, Application US/09650855
; Patent No. 6365355
; GENERAL INFORMATION:
; APPLICANT: MCCUTHEN-MALONEY, SANDRA
; APPLICANT: LAWRENCE LIVERMORE NATIONAL LABORATORY
; TITLE OF INVENTION: CHIMERIC PROTEINS FOR DETECTION AND QUANTITATION OF DNA
; TITLE OF INVENTION: MUTATIONS, DNA SEQUENCE VARIATIONS, DNA DAMAGE AND DNA
; FILE REFERENCE: IL-10284
; CURRENT APPLICATION NUMBER: US/09/650,855
; CURRENT FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/192,764
; PRIOR FILING DATE: 2000-03-28
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 101
; LENGTH: 940
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-650-855-101

Query Match
Best Local Similarity 33.3%; Score 5; DB 4; Length 940;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRA 14
Db 719 PESRA 723

RESULT 169
US-08-006-676B-1
; Sequence 1, Application US/0800676B
; Patent No. 5411865
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven
; TITLE OF INVENTION: Diagnosis of Leishmaniasis
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Jeffrey B. Oster
; STREET: 8339 SE 57th Street
; CITY: Mercer Island
; STATE: Washington
; COUNTRY: USA
; ZIP: 98040-4906
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORD for Windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/006,676B
FILING DATE: 15-JAN-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oster, Jeffrey B.
REGISTRATION NUMBER: 32,585
REFERENCE/DOCKET NUMBER: REED-4
TELEPHONE: (206) 232 7845
TELEFAX: (206) 236 0205
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 955 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-006-676B-1

Query Match 33.3%; Score 5; DB 1; Length 955;
Best Local Similarity 100.0%; Pred. No. 8.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 ESRAA 15
Db 478 ESRAA 482

RESULT 170
US-08-282-845-2
Sequence 2, Application US/08282845
Patent No. 5719263
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
TITLE OF INVENTION: A 230Kd Antigen Present in Leishmania
TITLE OF INVENTION: Species
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple Macintosh Operating System 7.1
SOFTWARE: Microsoft Word for Macintosh 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/282,845
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/006,676
FILING DATE: JANUARY 15, 1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: 5004-A
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 955 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-282-845-2

Query Match 33.3%; Score 5; DB 1; Length 955;
Best Local Similarity 100.0%; Pred. No. 8.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 ESRAA 15
Db 478 ESRAA 482

RESULT 171
US-08-428-414A-3
Sequence 3, Application US/08428414A
Patent No. 5912166
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
TITLE OF INVENTION: LEISHMANIASIS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/428,414A
FILING DATE: 21-APR-1995
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Radlecek, Ann T.
REGISTRATION NUMBER: 39,244
REFERENCE/DOCKET NUMBER: 210121.407
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
TELEX: 3723836 SEDANDBERRY
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 955 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-428-414A-3

Query Match 33.3%; Score 5; DB 2; Length 955;
Best Local Similarity 100.0%; Pred. No. 8.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 ESRAA 15
Db 478 ESRAA 482

RESULT 172
PCT-US94-00324-1
Sequence 1, Application PC/TUS9400324
GENERAL INFORMATION:
APPLICANT: Reed, Steven
TITLE OF INVENTION: Diagnosis of Leishmaniasis
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: Washington

COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple System 7.1
SOFTWARE: Microsoft Word, version 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/00324
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/006,676
FILING DATE: 15-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: 5004-MO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
INFORMATION FOR SEQ. ID NO. 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 955 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-00324-1

Query Match 33.3%; Score 5; DB 5; Length 955;
Best Local Similarity 100.0%; Pred. No. 8.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRA 15
DB 478 ESRA 482

RESULT 173
US-08-180-195-36
Sequence 36, Application US/08180195
Patent No. 5567584
GENERAL INFORMATION:
APPLICANT: Sledziewski Ph.D., Andrzej Z
APPLICANT: Bell, Lillian A.
APPLICANT: Kindvogel Ph.D., Wayne R.
TITLE OF INVENTION: METHODS OF PRODUCING SECRETED RECEPTOR ANALOGS
TITLE OF INVENTION: AND BIOLOGICALLY ACTIVE DIMERIZED POLYPEPTIDE
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry
STREET: 6300 Columbia Center
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/180,195
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/634,510
FILING DATE:
APPLICATION NUMBER: US 07/146,877
FILING DATE: 22-JAN-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/347,291

FILING DATE: 02-MAY-1989
ATTORNEY/AGENT INFORMATION:
NAME: Maki J.D., David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 990008.446C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031
TELEX: 3723836
INFORMATION FOR SEQ. ID NO. 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 1089 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-180-195-36

Query Match 33.3%; Score 5; DB 1; Length 1089;
Best Local Similarity 100.0%; Pred. No. 9.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HLGPH 6
DB 650 HLGPH 654

RESULT 174
US-08-168-917-4
Sequence 4, Application US/08168917
Patent No. 5686572
GENERAL INFORMATION:
APPLICANT: Wolf, David
APPLICANT: Tomlinson, James E.
APPLICANT: Fretto, Larry J.
APPLICANT: Giese, Neil A.
APPLICANT: Escobedo, Jaime A.
APPLICANT: Williams, Lewis T.
TITLE OF INVENTION: DOMAINS OF EXTRACELLULAR REGION OF HUMAN
TITLE OF INVENTION: PLATELET-DERIVED GROWTH FACTOR RECEPTOR POLYPEPTIDES
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND
STREET: Steuart Street Tower, 20th Floor \ One Market
STREET: Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/168,917
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/650,793
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: 12418-14
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ. ID NO. 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1089 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-168-917-4

Query Match

33.3%; Score 5; DB 1; Length 1089;
Best Local Similarity 100.0%; Pred. No. 9.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;QY 2 HLGPB 6
|||||

DB 650 HLGPB 654

RESULT 175

US-08-477-329-36
; Sequence 36, Application US/08477329
; Patent No. 5750375

GENERAL INFORMATION:

APPLICANT: Siedziewski Ph.D., Andrzej Z
APPLICANT: Bell, Lillian A.
TITLE OF INVENTION: METHODS OF PRODUCING SECRETED RECEPTOR ANALOGS AND BIOLOGICAL
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,329
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 990008.446C6
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031
TELEX: 3723836

INFORMATION FOR SEQ ID NO: 36:

SEQUENCE CHARACTERISTICS:
LENGTH: 1089 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-477-329-36

Query Match

33.3%; Score 5; DB 1; Length 1089;
Best Local Similarity 100.0%; Pred. No. 9.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;QY 2 HLGPB 6
|||||

DB 650 HLGPB 654

RESULT 176

US-08-475-458-36
; Sequence 36, Application US/08475458
; Patent No. 5843725

GENERAL INFORMATION:

APPLICANT: Siedziewski Ph.D., Andrzej Z
APPLICANT: Bell, Lillian A.
TITLE OF INVENTION: METHODS OF PRODUCING SECRETED RECEPTOR ANALOGS AND BIOLOGICAL
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:ADDRESSEE: Seed and Berry
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,458
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 990008.446D5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031
TELEX: 3723836

INFORMATION FOR SEQ ID NO: 36:

SEQUENCE CHARACTERISTICS:
LENGTH: 1089 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-475-458-36

Query Match

33.3%; Score 5; DB 2; Length 1089;
Best Local Similarity 100.0%; Pred. No. 9.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;QY 2 HLGPB 6
|||||

DB 650 HLGPB 654

RESULT 177

US-08-460-510-4
; Sequence 4, Application US/08460510
; Patent No. 5872218

GENERAL INFORMATION:

APPLICANT: Wolf, David
APPLICANT: Tomlinson, James E.
APPLICANT: Pretto, Larry J.
APPLICANT: Giese, Neil A.
APPLICANT: Escobedo, Jaime A.
APPLICANT: Williams, Lewis T.
TITLE OF INVENTION: DOMAINS OF EXTRACELLULAR REGION OF HUMAN
TITLE OF INVENTION: PLATELET-DERIVED GROWTH FACTOR RECEPTOR POLYPEPTIDES
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND and CREW
STREET: One Market Plaza, Stewart Street Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,510
FILING DATE: 02-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684

REFERENCE/DOCKET NUMBER: 012418-001430
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1089 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-460-510-4

Query Match 33.3%; Score 5; DB 2; Length 1089;
Best Local Similarity 100.0%; Pred. No. 9.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HLGP 6
|||||
DB 650 HLGP 654

RESULT 178
US-08-460-490-4
Sequence 4, Application US/08460490
Patent No. 5891652
GENERAL INFORMATION:
APPLICANT: Wolf, David
APPLICANT: Tomlinson, James E.
APPLICANT: Fretto, Larry J.
APPLICANT: Giese, Neill A.
APPLICANT: Escobedo, Jaime A.
APPLICANT: Williams, Lewis T.
TITLE OF INVENTION: DOMAINS OF EXTRACELLULAR REGION OF HUMAN
NUMBER OF SEQUENCES: 23
TITLE OF INVENTION: PLATELET-DERIVED GROWTH FACTOR RECEPTOR POLYPEPTIDES
CORRESPONDENCE ADDRESS:
ADDRESS: TOWNSEND and TOWNSEND and CREW
STREET: One Market Plaza, Stewart Street Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,490
FILING DATE: 02-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 012418-001420
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1089 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-460-490-4

Query Match 33.3%; Score 5; DB 2; Length 1089;
Best Local Similarity 100.0%; Pred. No. 9.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 HLGP 6
|||||
DB 650 HLGP 654

RESULT 179
US-08-980-400-36
Sequence 36, Application US/08980400
Patent No. 6018026
GENERAL INFORMATION:
APPLICANT: Sledziewski Ph.D., Andrzej Z
APPLICANT: Bell, Lillian A.
APPLICANT: Kinsvogel Ph.D., Wayne R.
TITLE OF INVENTION: METHODS OF PRODUCING SECRETED RECEPTOR ANALOGS AND BIOLOGIC
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESS: Seed and Berry
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/980,400
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/477,329
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Marki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 990008.446C6
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031
TELEX: 3723836
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 1089 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-980-400-36

Query Match 33.3%; Score 5; DB 3; Length 1089;
Best Local Similarity 100.0%; Pred. No. 9.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HLGP 6
|||||
DB 650 HLGP 654

RESULT 180
US-08-462-728-2
Sequence 2, Application US/08462728
Patent No. 6043211
GENERAL INFORMATION:
APPLICANT: WILLIAMS, Lewis T.
APPLICANT: ESCOBEDO, Jaime A.
TITLE OF INVENTION: PLATELET-DERIVED GROWTH FACTOR RECEPTORS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend and Crew
STREET: One Market, Stewart Street Tower, 20th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,728
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/309,322
FILING DATE: 10-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/151,414
FILING DATE: 02-FEB-1988
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 2307K-267-2-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/326-2400
TELEFAX: 415/326-2422
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1089 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-462-728-2

Query Match 33.3%: Score 5; DB 3; Length 1089;
Best Local Similarity 100.0%; Pred. No. 9.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HLGP 6
|||||
Db 650 HLGP 654

RESULT 181
US-09-583-459A-36
Sequence 36, Application US/09583459A
Patent No. 6291212
GENERAL INFORMATION:
APPLICANT: Siedzlewski Ph.D., Andrzej Z
APPLICANT: Bell, Lillian A.
TITLE OF INVENTION: METHODS OF PRODUCING SECRETED RECEPTOR ANALOGS
TITLE OF INVENTION: AND BIOLOGICALLY ACTIVE DIMERIZED POLYPEPTIDE
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry
STREET: 6300 Columbia Center
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/583,459A
FILING DATE: 30-MAY-2000
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/634,510
FILING DATE: 27-DEC-1990
APPLICATION NUMBER: US 07/146,877
FILING DATE: 22-JAN-1988
APPLICATION NUMBER: US 07/347,291
FILING DATE: 02-MAY-1989

ATTORNEY/AGENT INFORMATION:
NAME: Maki J.D., David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 990008.446C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 1089 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-583-459A-36

Query Match 33.3%: Score 5; DB 4; Length 1089;
Best Local Similarity 100.0%; Pred. No. 9.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HLGP 6
|||||
Db 650 HLGP 654

RESULT 182
US-09-583-210-36
Sequence 36, Application US/09583210
Patent No. 6291646
GENERAL INFORMATION:
APPLICANT: Siedzlewski Ph.D., Andrzej Z
APPLICANT: Bell, Lillian A.
TITLE OF INVENTION: METHODS OF PRODUCING SECRETED RECEPTOR ANALOGS
TITLE OF INVENTION: AND BIOLOGICALLY ACTIVE DIMERIZED POLYPEPTIDE
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry
STREET: 6300 Columbia Center
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/583,210
FILING DATE: 30-MAY-2000
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/634,510
FILING DATE: 27-DEC-1990
APPLICATION NUMBER: US 07/146,877
FILING DATE: 22-JAN-1988
APPLICATION NUMBER: US 07/347,291
FILING DATE: 02-MAY-1989
ATTORNEY/AGENT INFORMATION:
NAME: Maki J.D., David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 990008.446C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 1089 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
US-09-583-210-36

Query Match 33.3%; Score 5; DB 4; Length 1089;
Best Local Similarity 100.0%; Pred. No. 9.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HLGPB 6
|||||
Db 650 HLGPB 654

RESULT 183

US-09-583-449A-36
Sequence 36, Application US/09583449A
Patent No. 6300099

GENERAL INFORMATION:

APPLICANT: Siedzielski Ph.D., Andrzej Z
APPLICANT: Bell, Lillian A.
APPLICANT: Kindsvogel Ph.D., Wayne R.
TITLE OF INVENTION: METHODS OF PRODUCING SECRETED RECEPTOR ANALOGS
TITLE OF INVENTION: AND BIOLOGICALLY ACTIVE DIMERIZED POLYPEPTIDE
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry
STREET: 6300 Columbia Center
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/583,449A
FILING DATE: 30-MAY-2000
CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/634,510
FILING DATE: 27-DEC-1990
APPLICATION NUMBER: US 07/146,877
FILING DATE: 22-JAN-1988
APPLICATION NUMBER: US 07/347,291
FILING DATE: 02-MAY-1989

ATTORNEY/AGENT INFORMATION:

NAME: Maki J.D., David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 990008.446C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031

INFORMATION FOR SEQ. ID NO: 36:

SEQUENCE CHARACTERISTICS:
LENGTH: 1089 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-583-449A-36

Query Match 33.3%; Score 5; DB 4; Length 1089;

Best Local Similarity 100.0%; Pred. No. 9.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HLGPB 6
|||||
Db 650 HLGPB 654

RESULT 184

US-09-435-059-36
Sequence 36, Application US/09435059
Patent No. 6323323

GENERAL INFORMATION:
APPLICANT: Siedzielski Ph.D., Andrzej Z
APPLICANT: Bell, Lillian A.
APPLICANT: Kindsvogel Ph.D., Wayne R.
TITLE OF INVENTION: METHODS OF PRODUCING SECRETED RECEPTOR ANALOGS AND BIOLOGIC
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/435,059
FILING DATE:
CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/477,329
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 990008.446C6
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031

INFORMATION FOR SEQ. ID NO: 36:

SEQUENCE CHARACTERISTICS:
LENGTH: 1089 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-435-059-36

Query Match 33.3%; Score 5; DB 4; Length 1089;

Best Local Similarity 100.0%; Pred. No. 9.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HLGPB 6
|||||
Db 650 HLGPB 654

RESULT 185

US-08-461-917-2
Sequence 2, Application US/08461917
Patent No. 6372438

GENERAL INFORMATION:

APPLICANT: WILLIAMS, Lewis T.
APPLICANT: ESCOBEDO, Jaime A.
TITLE OF INVENTION: PLATELET-DERIVED GROWTH FACTOR RECEPTORS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew
STREET: One Market, Stewart Street Tower, 20th Floor
CITY: San Francisco
STATE: California

COUNTRY: USA
ZIP: 94105

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,917
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/309,322
FILING DATE: 10-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/151,414
FILING DATE: 02-FEB-1988
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 2307K-267-2-4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/326-2400
TELEFAX: 415/326-2422
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1089 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-461-917-2

Query Match 33.3%; Score 5; DB 4; Length 1089;
Best Local Similarity 100.0%; Pred. No. 9.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HLGP 6
|||||
Db 650 HLGP 654

RESULT 186
PCT-US92-00730-4
Sequence 4, Application PC/TUS9200730
GENERAL INFORMATION:
APPLICANT: Wolf, David
APPLICANT: Tomlinson, James E.
APPLICANT: Fretto, Larry J.
APPLICANT: Giese, Neil A.
APPLICANT: Escobedo, Jaime A.
APPLICANT: Williams, Jaime T.
TITLE OF INVENTION: DOMAINS OF EXTRACELLULAR REGION OF HUMAN
TITLE OF INVENTION: PLATELET-DERIVED GROWTH FACTOR RECEPTOR POLYPEPTIDES
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND
STREET: Steuart Street Tower, 20th Floor \ One Market
CITY: Plaza
STATE: San Francisco
COUNTRY: US
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/00730
FILING DATE: 19920128
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: 12418-14
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
LENGTH: 1089 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US92-00730-4

Query Match 33.3%; Score 5; DB 5; Length 1089;
Best Local Similarity 100.0%; Pred. No. 9.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HLGP 6
|||||
Db 650 HLGP 654

RESULT 187
PCT-US92-00862-4
Sequence 4, Application PC/TUS9200862
GENERAL INFORMATION:
APPLICANT: Williams, Lewis T.
APPLICANT: Escobedo, Jaime A.
TITLE OF INVENTION: PLATELET-DERIVED GROWTH FACTOR RECEPTORS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND
STREET: Steuart Street Tower, 20th Floor
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/00862
FILING DATE: 19920131
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/151,141
FILING DATE: 02-FEB-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/309,322
FILING DATE: 10-FEB-1989
ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: 2307U-267-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1089 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US92-00862-4

Query Match 33.3%; Score 5; DB 5; Length 1089;
Best Local Similarity 100.0%; Pred. No. 9.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HLGP 6
|||||
Db 650 HLGP 654

RESULT 188
US-08-574-959A-7
Sequence 7, Application US/08574959A

Patent No. 5962224
GENERAL INFORMATION:
APPLICANT: Jaekyoon Shin, Insil Joungh, Ratna K. Vadlamudi
APPLICANT: and Jack L. Strominger
TITLE OF INVENTION: P62 POLYPEPTIDES, RELATED POLYPEPTIDES
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHYE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/574,959A
FILING DATE: 19-DEC-95
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: DFN-008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1135 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-574-959A-7

Query Match 33.3%; Score 5; DB 2; Length 1135;
Best Local Similarity 100.0%; Pred. No. 9.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GPHRS 8
Db 659 GPHRS 663

RESULT 189
US-09-357-014-7
Sequence 7, Application US/09357014
Patent No. 6291645
GENERAL INFORMATION:
APPLICANT: Jaekyoon Shin, Insil Joungh, Ratna K. Vadlamudi
APPLICANT: and Jack L. Strominger
TITLE OF INVENTION: P62 POLYPEPTIDES, RELATED POLYPEPTIDES,
AND USES THEREFOR
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHYE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/357,014
FILING DATE: 19-Jul-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/574,959

FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: DFN-008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1135 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-357-014-7

Query Match 33.3%; Score 5; DB 4; Length 1135;
Best Local Similarity 100.0%; Pred. No. 9.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GPHRS 8
Db 659 GPHRS 663

RESULT 190
US-08-323-474-8
Sequence 8, Application US/08323474
Patent No. 5447860
GENERAL INFORMATION:
APPLICANT: Ziegler, Steven F.
TITLE OF INVENTION: NOVEL TYROSINE KINASE
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/323,474
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/905,600
FILING DATE: 26-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Seese, Kathryn A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2609
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
TELEX: 756822
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1138 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-323-474-8

Query Match 33.3%; Score 5; DB 1; Length 1138;
Best Local Similarity 100.0%; Pred. No. 9.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRAA 15
11111
Db 749 ESRAA 753

RESULT 191

US-08-469-537A-98
; Sequence 98, Application US/08469537A
; Patent No. 5843749
; GENERAL INFORMATION:
; APPLICANT: Maisonneville, et al.
; TITLE OF INVENTION: EHK AND ROR TYROSINE
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.
; STREET: 777 Old Saw Mill River Road
; CITY: Tarrytown
; STATE: NY
; COUNTRY: U.S.A.
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,537A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/406,247
; FILING DATE: 17-MAR-1995
; APPLICATION NUMBER: USSN 08/144,992
; FILING DATE: 28-OCT-1993
; APPLICATION NUMBER: USSN 07/736,559
; FILING DATE: 26-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Kempler, Ph.D., Gal1 M
; REGISTRATION NUMBER: 32,143
; REFERENCE/DOCKET NUMBER: REG 070C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 914-345-7400
; TELEFAX: 914-345-7721
; TELEX:
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1138 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-469-537A-98

Query Match 33.3%; Score 5; DB 2; Length 1138;
Best Local Similarity 100.0%; Pred. No. 9.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRAA 15
11111
Db 749 ESRAA 753

RESULT 192

US-08-220-240A-5
; Sequence 5, Application US/08220240A
; Patent No. 5955291
; GENERAL INFORMATION:
; APPLICANT: Aitalo, Kari
; APPLICANT: Mälikäinen, Marja-Terttu
; APPLICANT: Partanen, Juha
; APPLICANT: Mäkelä, Tomi

APPLICANT: Korhonen, Jaana
; TITLE OF INVENTION: ANTIBODIES RECOGNIZING TIE RECEPTOR
; TITLE OF INVENTION: TYROSINE KINASE AND USBS THEREOF
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/220,240A
; FILING DATE: 29-MAR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FI93/00006
; FILING DATE: 08-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/817,800
; FILING DATE: 09-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/167,453
; FILING DATE: 15-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Gass, David A.
; REGISTRATION NUMBER: 38,153
; REFERENCE/DOCKET NUMBER: 29151/31958
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-220-240A-5

Query Match 33.3%; Score 5; DB 2; Length 1138;
Best Local Similarity 100.0%; Pred. No. 9.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRAA 15
11111
Db 749 ESRAA 753

RESULT 193

PCT-US93-03076-2
; Sequence 2, Application PC/TUS9303076
; GENERAL INFORMATION:
; APPLICANT: Whitehead Institute for Biomedical Research
; TITLE OF INVENTION: GAP-Associated Protein p190 and
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: 2 Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US93/03076
;; FILING DATE: 19930331
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Granahan, Patricia
;; REGISTRATION NUMBER: 32,227
;; REFERENCE/DOCKET NUMBER: WH192-03A
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 617-861-6240
;; TELEFAX: 617-861-9540
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1513 amino acids
;; TYPE: AMINO ACID
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
PCT-US93-03076-2

Query Match 33.3%; Score 5; DB 5; Length 1513;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 PHRS 9
Db 107 PHRS 111

RESULT 194
US-08-479-722B-2
; Sequence 2, Application US/08479722B
; Patent No. 6074840
; GENERAL INFORMATION:
; APPLICANT: Bonadio, Jeffrey
; APPLICANT: Yin, Wushan
; TITLE OF INVENTION: LATENT TGF(BINDING PROTEIN (LTBP)
; TITLE OF INVENTION: GENES, COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Williams, Morgan & Amereson
; STREET: 7676 Hillmont, Suite 250
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77040
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,722B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US95/02251
; FILING DATE: 21-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/316,650
; FILING DATE: 30-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/199,780
; FILING DATE: 18-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Fussey, Shelley P.M.
; REGISTRATION NUMBER: 39,458
; REFERENCE/DOCKET NUMBER: 4100.000500/FUS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 934-7000
; TELEFAX: (713) 934-7011
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1833 amino acids

;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-479-722B-2

Query Match 33.3%; Score 5; DB 3; Length 1833;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 GPHRS 8
Db 259 GPHRS 263

RESULT 195
PCT-US95-02251-18
; Sequence 18, Application PC/TUS9502251
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STIMULATING BONE
; TITLE OF INVENTION: CELLS
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; SOFTWARE: PatentIn Release #1.0, Version
; SOFTWARE: #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02251
; FILING DATE: CONCURRENTLY HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/316,650
; FILING DATE: 30-SEP-1994
; CLASSIFICATION:
; APPLICATION NUMBER: US 08/199,780
; FILING DATE: 18-FEB-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UMIC009P--
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEEX: 79-0924
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1833 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-02251-18

Query Match 33.3%; Score 5; DB 5; Length 1833;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 GPHRS 8
Db 259 GPHRS 263

RESULT 196
US-09-091-219-24

; Sequence 24, Application US/09091219
; Patent No. 6171592
; GENERAL INFORMATION:
; APPLICANT: STUDDERT, Michael J.
; APPLICANT: CRABB, Brendan S.
; APPLICANT: FENG, Li
; TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS
; FILE REFERENCE: 040268/0151
; CURRENT APPLICATION NUMBER: US/09/091,219
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: PCT/AU96/00815
; EARLIER FILING DATE: 1996-12-18
; EARLIER APPLICATION NUMBER: AU PN7201
; EARLIER FILING DATE: 1995-12-18
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 2318
; TYPE: PRT
; ORGANISM: Foot-and-mouth disease virus
; US-09-091-219-24

Query Match 33.3%; Score 5; DB 4; Length 2318;
Best Local Similarity 100.0%; Pred. No. 1.8e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
|||||
DB 1095 RSTPE 1099

RESULT 197
US-08-602-999A-97
; Sequence 97, Application US/08602999A
; Patent No. 6184205
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: KAY, Brian K.
; APPLICANT: THORN, Judith M.
; APPLICANT: OUILIAM, Lawrence A.
; APPLICANT: DER, Channing J.
; APPLICANT: FOWLER, Dana M.
; APPLICANT: RIDER, James E.
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
; TITLE OF INVENTION: ISOLATING AND USING SAME
; NUMBER OF SEQUENCES: 467
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/602,999A
; FILING DATE: 16-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MISTOCK, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-202
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; INFORMATION FOR SEQ ID NO: 97:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids

; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-602-999A-97

Query Match 26.7%; Score 4; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 10
|||||
DB 1 RSTPE 4

RESULT 198
US-08-496-841C-162
; Sequence 162, Application US/08496841C
; Patent No. 6210919
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; ROMKENS, JOHANNA M
; FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 175
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Darby & Darby, PC
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/496,841C
; FILING DATE: 28-Jun-1995
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul F. Fehlner, Ph.D.
; REGISTRATION NUMBER: 35,135
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 527-7700
; TELEFAX: (212) 753-6237
; INFORMATION FOR SEQ ID NO: 162:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 162:
US-08-496-841C-162

Query Match 26.7%; Score 4; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 RSTPE 11
|||||
DB 1 RSTPE 4

RESULT 199
US-08-278-865-97
; Sequence 97, Application US/08278865
; Patent No. 6303574
; GENERAL INFORMATION:
; APPLICANT: KAY, BRIAN K.
; APPLICANT: SPARKS, ANDREW B.

APPLICANT: THORN, JUDITH M.
APPLICANT: OULLIAM, LAWRENCE A.
APPLICANT: DER, CHANNING J.
TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
TITLE OF INVENTION: ISOLATING AND USING SAME
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/278,865
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Villacorta, Gilberto M.
REGISTRATION NUMBER: 34,038
REFERENCE/DOCKET NUMBER: 4980-007-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 97:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-278-865-97

Query Match 26.7%; Score 4; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTP 10
|||||
Db 1 RSTP 4

RESULT 200
US-09-500-124-97
Sequence 97, Application US/09500124
Patent No. 6432920
GENERAL INFORMATION:
APPLICANT: SPARKS, Andrew B.
APPLICANT: KAY, Brian K.
APPLICANT: THORN, Judith M.
APPLICANT: OULLIAM, Lawrence A.
APPLICANT: DER, Channing J.
APPLICANT: FOWLKES, Dana M.
APPLICANT: RIDER, James E.
TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
TITLE OF INVENTION: ISOLATING AND USING SAME
NUMBER OF SEQUENCES: 467
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/500,124
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/602,999
FILING DATE: 16-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: MISTOCK, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-202
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 97:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-09-500-124-97

Query Match 26.7%; Score 4; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTP 10
|||||
Db 1 RSTP 4

Search completed: March 10, 2003, 14:31:17
job time : 38 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2003, 14:30:05 ; Search time 89 Seconds

(without alignments)
7.107 Million cell updates/sec

Title: US-09-689-159A-2_COPY_346_360

Perfect score: 15
Sequence: 1 SHLPHRSTPESRAA 15

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 188354 seqs, 42170167 residues

Word size : 0

Total number of hits satisfying chosen parameters: 188354

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 200 summaries

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15	100.0	354	10	US-09-823-153-8
2	15	100.0	463	10	US-09-895-035-12
3	15	100.0	467	10	US-09-785-474-2
4	15	100.0	467	10	US-09-785-474-4
5	15	100.0	467	10	US-09-785-474-28
6	15	100.0	467	10	US-09-785-474-32
7	15	100.0	467	10	US-09-785-474-30
8	15	100.0	467	10	US-09-754-949-4
9	15	100.0	467	10	US-09-878-454A-3
10	15	100.0	467	10	US-09-895-035-14
11	15	100.0	467	12	US-10-071-900-1
12	13	86.7	15	10	US-09-823-153-9
13	13	86.7	16	10	US-09-823-153-5
14	6	40.0	19	10	US-09-734-520-16
15	6	40.0	19	12	US-10-012-034A-16
16	6	40.0	55	10	US-09-864-761-38777
17	6	40.0	429	10	US-09-815-242-15506
18	6	40.0	739	9	US-09-738-626-6773
19	6	40.0	1086	10	US-09-147-052-4

20	6	40.0	1090	10	US-09-866-510-14	Sequence 14, Appl
21	6	40.0	1106	9	US-09-955-363-2	Sequence 2, Appl
22	6	40.0	1106	10	US-09-866-510-16	Sequence 16, Appl
23	6	40.0	1106	10	US-09-866-510-18	Sequence 18, Appl
24	6	40.0	1106	10	US-09-866-510-20	Sequence 20, Appl
25	6	40.0	1106	10	US-09-866-510-22	Sequence 22, Appl
26	6	40.0	1106	10	US-09-734-520-75	Sequence 75, Appl
27	5	33.3	12	10	US-09-734-520-75	Sequence 75, Appl
28	5	33.3	12	12	US-10-012-034A-75	Sequence 75, Appl
29	5	33.3	12	12	US-10-012-034A-76	Sequence 76, Appl
30	5	33.3	19	10	US-09-734-520-15	Sequence 15, Appl
31	5	33.3	19	12	US-10-012-034A-15	Sequence 15, Appl
32	5	33.3	34	10	US-09-864-761-44095	Sequence 44095, A
33	5	33.3	39	9	US-09-774-639-274	Sequence 274, App
34	5	33.3	39	9	US-09-774-639-351	Sequence 351, App
35	5	33.3	48	10	US-09-864-761-44670	Sequence 44670, A
36	5	33.3	91	10	US-09-764-847-695	Sequence 695, App
37	5	33.3	91	10	US-09-864-761-33764	Sequence 33764, A
38	5	33.3	97	9	US-09-738-626-5971	Sequence 5971, Ap
39	5	33.3	97	10	US-09-852-137-2	Sequence 2, Appl
40	5	33.3	98	10	US-09-864-761-41798	Sequence 41798, A
41	5	33.3	109	9	US-09-736-457-1671	Sequence 1671, Ap
42	5	33.3	109	9	US-09-902-941-1671	Sequence 1671, Ap
43	5	33.3	109	9	US-09-849-626-1671	Sequence 1671, Ap
44	5	33.3	114	10	US-09-925-300-1175	Sequence 1175, Ap
45	5	33.3	121	9	US-10-004-551-12	Sequence 12, Appl
46	5	33.3	126	9	US-09-738-626-6171	Sequence 6171, Ap
47	5	33.3	126	10	US-09-734-559-86	Sequence 86, Appl
48	5	33.3	129	9	US-09-764-868-781	Sequence 781, App
49	5	33.3	130	9	US-09-738-626-6788	Sequence 6788, Ap
50	5	33.3	134	10	US-09-764-864-1425	Sequence 1425, Ap
51	5	33.3	146	9	US-09-764-868-1017	Sequence 1017, Ap
52	5	33.3	149	9	US-09-738-626-4628	Sequence 4628, Ap
53	5	33.3	161	10	US-09-747-155-243	Sequence 243, App
54	5	33.3	164	9	US-09-736-457-798	Sequence 798, App
55	5	33.3	164	9	US-09-902-941-798	Sequence 798, App
56	5	33.3	164	9	US-09-849-626-798	Sequence 798, App
57	5	33.3	168	10	US-09-864-761-48050	Sequence 48050, A
58	5	33.3	177	9	US-09-736-457-1678	Sequence 1678, Ap
59	5	33.3	177	9	US-09-902-941-1678	Sequence 1678, Ap
60	5	33.3	181	9	US-09-849-626-1678	Sequence 1678, Ap
61	5	33.3	181	9	US-10-093-766-52	Sequence 52, Appl
62	5	33.3	186	10	US-09-811-284-176	Sequence 176, App
63	5	33.3	192	9	US-09-996-015-43	Sequence 43, Appl
64	5	33.3	193	9	US-09-996-015-44	Sequence 44, Appl
65	5	33.3	193	9	US-09-996-015-47	Sequence 47, Appl
66	5	33.3	202	9	US-09-996-015-8	Sequence 8, Appl
67	5	33.3	205	9	US-09-738-626-5128	Sequence 5128, App
68	5	33.3	232	10	US-09-925-301-898	Sequence 898, App
69	5	33.3	248	10	US-09-815-242-5905	Sequence 5905, Ap
70	5	33.3	269	10	US-09-764-865-658	Sequence 658, App
71	5	33.3	271	9	US-09-908-193-40	Sequence 40, Appl
72	5	33.3	280	9	US-09-738-626-4071	Sequence 4071, Ap
73	5	33.3	310	10	US-09-925-300-1344	Sequence 1344, Ap
74	5	33.3	311	9	US-09-764-868-1009	Sequence 1009, Ap
75	5	33.3	311	9	US-09-791-932-63	Sequence 63, Appl
76	5	33.3	314	9	US-09-736-457-1863	Sequence 1863, Ap
77	5	33.3	314	9	US-09-902-941-1863	Sequence 1863, Ap
78	5	33.3	318	9	US-09-849-626-1863	Sequence 1863, Ap
79	5	33.3	318	9	US-09-975-719-23	Sequence 23, Appl
80	5	33.3	327	9	US-09-908-193-39	Sequence 39, Appl
81	5	33.3	329	9	US-09-738-626-4330	Sequence 4350, Ap
82	5	33.3	329	10	US-09-886-055-211	Sequence 211, App
83	5	33.3	335	9	US-09-908-193-37	Sequence 37, Appl
84	5	33.3	336	9	US-09-908-193-38	Sequence 38, Appl
85	5	33.3	338	10	US-09-815-242-12043	Sequence 12043, A
86	5	33.3	348	9	US-10-114-170-263	Sequence 263, App
87	5	33.3	350	10	US-09-908-805B-25	Sequence 25, Appl
88	5	33.3	358	12	US-10-108-714-4	Sequence 4, Appl
89	5	33.3	360	9	US-09-738-626-3763	Sequence 3763, Ap
90	5	33.3	365	9	US-09-870-759-73	Sequence 73, Appl
91	5	33.3	380	9	US-09-165-522-16	Sequence 16, Appl
92	5	33.3	386	9	US-09-939-833-6	Sequence 6, Appl

Db 291 SHLGPHRSTPESRAA 305

RESULT 2

US-09-895-035-12
Sequence 12, Application US/09895035
Patent No. US20020082211A1
GENERAL INFORMATION:
APPLICANT: Patterson, Chandra
APPLICANT: Murry, Lynn E.
APPLICANT: Kaser, Matthew R.
TITLE OF INVENTION: HUMAN PRESENTILIN VARIANT
FILE REFERENCE: PC-0047 CIP
CURRENT FILING DATE: 2001-06-29
PRIOR FILING DATE: 1998-07-16
PRIOR APPLICATION NUMBER: 09/116,640
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PERL Program
SEQ ID NO 12
LENGTH: 463
TYPE: PRP
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No. US20020082211A1 g1244638
US-09-895-035-12

Query Match 100.0%; Score 15; DB 10; Length 463;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPHRSTPESRAA 15
Db 342 SHLGPHRSTPESRAA 356

RESULT 3

US-09-785-474-2
Sequence 2, Application US/09785474
Patent No. US20010012626A1
GENERAL INFORMATION:
APPLICANT: TANZI, RUDOLPH
APPLICANT: MASCO, WILMA
TITLE OF INVENTION: Genetic Alterations Related To Familial
Alzheimer's Disease
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVENUE, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/785,474
FILING DATE: 20-Feb-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/706,344
FILING DATE: 30-AUG-1996
APPLICATION NUMBER: 60/003,054
FILING DATE: 31-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: KIM, JUDITH U.
REGISTRATION NUMBER: 40,679
REFERENCE/DOCKET NUMBER: 0609.4180002

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 467 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-785-474-2

Query Match 100.0%; Score 15; DB 10; Length 467;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPHRSTPESRAA 15
Db 346 SHLGPHRSTPESRAA 360

RESULT 4

US-09-785-474-4
Sequence 4, Application US/09785474
Patent No. US20010012626A1
GENERAL INFORMATION:
APPLICANT: TANZI, RUDOLPH
APPLICANT: MASCO, WILMA
TITLE OF INVENTION: Genetic Alterations Related To Familial
Alzheimer's Disease
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVENUE, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/785,474
FILING DATE: 20-Feb-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/706,344
FILING DATE: 30-AUG-1996
APPLICATION NUMBER: 60/003,054
FILING DATE: 31-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: KIM, JUDITH U.
REGISTRATION NUMBER: 40,679
REFERENCE/DOCKET NUMBER: 0609.4180002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-785-474-4

Query Match 100.0%; Score 15; DB 10; Length 467;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPHRSTPESRAA 15

Db 346 SHLGPBRSPTESRAA 360

RESULT 5
US-09-785-474-28

Sequence 28, Application US/09785474
Patent No. US20010012626A1

GENERAL INFORMATION:

APPLICANT: TANZI, RUDOLPH

MASCO, WILMA

TITLE OF INVENTION: Genetic Alterations Related To Familial Alzheimer's Disease

NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESSES:

ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.

STREET: 1100 NEW YORK AVENUE, SUITE 600

CITY: WASHINGTON

STATE: DC

COUNTRY: USA

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/785,474

FILING DATE: 20-Feb-2001

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/706,344

FILING DATE: 30-AUG-1996

APPLICATION NUMBER: 60/003,054

FILING DATE: 31-AUG-1995

ATTORNEY/AGENT INFORMATION:

NAME: KIM, JUDITH U.

REGISTRATION NUMBER: 40,679

REFERENCE/DOCKET NUMBER: 0609,4180002

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2600

TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 28:

SEQUENCE CHARACTERISTICS:

LENGTH: 467 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 28:

US-09-785-474-28

Query Match 100.0%; Score 15; DB 10; Length 467;

Best Local Similarity 100.0%; Pred. No. 1,9e-08;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SHLGPBRSPTESRAA 15

Db 346 SHLGPBRSPTESRAA 360

RESULT 6
US-09-785-474-30

Sequence 30, Application US/09785474

Patent No. US20010012626A1

GENERAL INFORMATION:

APPLICANT: TANZI, RUDOLPH

MASCO, WILMA

TITLE OF INVENTION: Genetic Alterations Related To Familial Alzheimer's Disease

NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESSES:

ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.

STREET: 1100 NEW YORK AVENUE, SUITE 600

CITY: WASHINGTON

STATE: DC

COUNTRY: USA

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/785,474

FILING DATE: 20-Feb-2001

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/706,344

FILING DATE: 30-AUG-1996

APPLICATION NUMBER: 60/003,054

FILING DATE: 31-AUG-1995

ATTORNEY/AGENT INFORMATION:

NAME: KIM, JUDITH U.

REGISTRATION NUMBER: 40,679

REFERENCE/DOCKET NUMBER: 0609,4180002

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2600

TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 30:

SEQUENCE CHARACTERISTICS:

LENGTH: 467 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 30:

US-09-785-474-30

Query Match 100.0%; Score 15; DB 10; Length 467;

Best Local Similarity 100.0%; Pred. No. 1,9e-08;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SHLGPBRSPTESRAA 15

Db 346 SHLGPBRSPTESRAA 360

RESULT 7
US-09-785-474-32

Sequence 32, Application US/09785474

Patent No. US20010012626A1

GENERAL INFORMATION:

APPLICANT: TANZI, RUDOLPH

MASCO, WILMA

TITLE OF INVENTION: Genetic Alterations Related To Familial Alzheimer's Disease

NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESSES:

ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.

STREET: 1100 NEW YORK AVENUE, SUITE 600

CITY: WASHINGTON

STATE: DC

COUNTRY: USA

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/785,474

FILING DATE: 20-Feb-2001

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/706,344

FILING DATE: 30-AUG-1996

APPLICATION NUMBER: 60/003,054

```
; FILING DATE: 31-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: KIM, JUDITH U.
; REGISTRATION NUMBER: 40,679
; REFERENCE/DOCKET NUMBER: 0609.4180002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 32:
US-09-785-474-32

Query Match          100.0%; Score 15; DB 10; Length 467;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SHLGPHRSTPESRAA 15
        |||
DB      346 SHLGPHRSTPESRAA 360

RESULT 8
US-09-754-949-4
; Sequence 4, Application US/09754949
; Patent No. US20020015939A1
; GENERAL INFORMATION:
; APPLICANT: MCCARTHY, JUSTIN
; APPLICANT: CORDELL, BARBARA
; TITLE OF INVENTION: METHODS FOR IDENTIFYING INHIBITORS OF
; FILE REFERENCE: SCIOS.012A
; CURRENT APPLICATION NUMBER: US/09/754,949
; CURRENT FILING DATE: 2001-01-04
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-754-949-4

Query Match          100.0%; Score 15; DB 10; Length 467;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SHLGPHRSTPESRAA 15
        |||
DB      346 SHLGPHRSTPESRAA 360

RESULT 9
US-09-878-454A-3
; Sequence 3, Application US/09878454A
; Patent No. US20020064828A1
; GENERAL INFORMATION:
; APPLICANT: Montelero, et al.
; TITLE OF INVENTION: Method of Controlling the Binding of Calmyrin to Presentlin
; FILE REFERENCE: 4115-161
; CURRENT APPLICATION NUMBER: US/09/878,454A
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 60/210,939
; PRIOR FILING DATE: 2000-06-11
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```
US-09-878-454A-3

Query Match          100.0%; Score 15; DB 10; Length 467;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SHLGPHRSTPESRAA 15
        |||
DB      346 SHLGPHRSTPESRAA 360

RESULT 10
US-09-895-035-14
; Sequence 14, Application US/09895035
; Patent No. US2002008211A1
; GENERAL INFORMATION:
; APPLICANT: Patterson, Chandra
; APPLICANT: Murry, Lynn E.
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: HUMAN PRESENTLIN VARIANT
; FILE REFERENCE: PC-0047 CIP
; CURRENT APPLICATION NUMBER: US/09/895,035
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 09/116,640
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PERL Program
; SEQ ID NO 14
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Inocyte ID No. US2002008211A1 g1709856
US-09-895-035-14

Query Match          100.0%; Score 15; DB 10; Length 467;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SHLGPHRSTPESRAA 15
        |||
DB      346 SHLGPHRSTPESRAA 360

RESULT 11
US-10-071-900-1
; Sequence 1, Application US/10071900
; Patent No. US20020127541A1
; GENERAL INFORMATION:
; APPLICANT: St. George-Hyslop, Peter H.
; APPLICANT: Rommens, Johanna
; APPLICANT: Fraser, Paul E.
; TITLE OF INVENTION: Alzheimer's Related Proteins and Methods
; FILE REFERENCE: 1034/1P810-US1
; CURRENT APPLICATION NUMBER: US/10/071,900
; CURRENT FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US/09/227,725
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-071-900-1

Query Match          100.0%; Score 15; DB 12; Length 467;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SHLGPHRSTPESRAA 15
```

Db 346 SHLGPBHRSTPESRAA 360

RESULT 12

US-09-823-153-9
; Sequence 9, Application US/09823153
; Patent No. US2002025540A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; APPLICANT: Roberts, Susan
; APPLICANT: Pak, Roger
; APPLICANT: Lewis, Martin
; APPLICANT: Smith, David
; APPLICANT: Hendrick, Joseph
; APPLICANT: Vinitsky, Alexander
; TITLE OF INVENTION: ISOLATION OF FUNCTIONALLY ACTIVE GAMMA-SECRETASE PROTEIN COMPLEX
; TITLE OF INVENTION: METHODS FOR DETECTION OF ACTIVITY AND INHIBITORS THEREOF
; FILE REFERENCE: D0004
; CURRENT APPLICATION NUMBER: US/09/823,153
; CURRENT FILING DATE: 2001-07-02
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Description of Artificial Sequence: PSI PEPTIDE
US-09-823-153-9

Query Match 86.7%; Score 13; DB 10; Length 15;
Best Local Similarity 100.0%; Pred. No. 1,1e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPBHRSTPESR 13
Db 3 SHLGPBHRSTPESR 15

RESULT 13

US-09-823-153-5
; Sequence 5, Application US/09823153
; Patent No. US20020025540A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; APPLICANT: Roberts, Susan
; APPLICANT: Pak, Roger
; APPLICANT: Lewis, Martin
; APPLICANT: Smith, David
; APPLICANT: Hendrick, Joseph
; APPLICANT: Vinitsky, Alexander
; TITLE OF INVENTION: ISOLATION OF FUNCTIONALLY ACTIVE GAMMA-SECRETASE PROTEIN COMPLEX
; TITLE OF INVENTION: METHODS FOR DETECTION OF ACTIVITY AND INHIBITORS THEREOF
; FILE REFERENCE: D0004
; CURRENT APPLICATION NUMBER: US/09/823,153
; CURRENT FILING DATE: 2001-07-02
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Description of Artificial Sequence: PSI PEPTIDE
US-09-823-153-5

Query Match 86.7%; Score 13; DB 10; Length 16;
Best Local Similarity 100.0%; Pred. No. 1,1e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPBHRSTPESR 13
Db 4 SHLGPBHRSTPESR 16

RESULT 14

US-09-734-520-16
; Sequence 16, Application US/09734520
; Patent No. US2002015173A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: SHORT PEPTIDES FROM THE A-REGION OF
; TITLE OF INVENTION: PROTEIN KINASES WHICH SELECTIVELY MODULATE PROTEIN KINASE
; TITLE OF INVENTION: ACTIVITY
; FILE REFERENCE: 1242 2003-000
; CURRENT APPLICATION NUMBER: US/09/734,520
; CURRENT FILING DATE: 2000-12-11
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: PDGFR-b
US-09-734-520-16

Query Match 40.0%; Score 6; DB 10; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPB 6
Db 5 SHLGPB 10

RESULT 15

US-10-012-034A-16
; Sequence 16, Application US/10012034A
; Patent No. US20020137141A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: SHORT PEPTIDES FROM THE A-REGION OF
; TITLE OF INVENTION: PROTEIN KINASES WHICH SELECTIVELY MODULATE PROTEIN KINASE
; TITLE OF INVENTION: ACTIVITY
; FILE REFERENCE: BEN-SASSON-5A
; CURRENT APPLICATION NUMBER: US/10/012,034A
; CURRENT FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 09/734,520
; PRIOR FILING DATE: 2000-12-11
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PDGFR-b
; FEATURE:
; NAME/KEY: MYRISTATE
; LOCATION: (1)...(0)
; FEATURE:
; NAME/KEY: AMIDATION
; LOCATION: (0)...(19)
US-10-012-034A-16

Query Match 40.0%; Score 6; DB 12; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPB 6
Db 5 SHLGPB 10


```
RESULT 16
US-09-864-761-38777
; Sequence 38777, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aemica-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 38777
; LENGTH: 55
; TYPE: prt
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007106.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 6.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 9.8
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 12
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 6.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 8.9
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.2
; OTHER INFORMATION: SWISSPROT HIT: P75320, EVALUATE 2.10e+00
US-09-864-761-38777

Query Match      40.0%; Score 6; DB 10; Length 55;
Best Local Similarity 100.0%; Pred. No. 5.9;
```

```
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 LGPHRS 8
Db 10 LGPHRS 15

RESULT 17
US-09-815-242-13506
; Sequence 13506, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13506
; LENGTH: 429
; TYPE: prt
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13506

Query Match      40.0%; Score 6; DB 10; Length 429;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 10 PESRAA 15
Db 171 PESRAA 176

RESULT 18
US-09-738-626-6773
; Sequence 6773, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, MAKINO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
```

```

; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO: 6773
; LENGTH: 739
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6773

Query Match
40.0%; Score 6; DB 9; Length 739;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRAA 15
DB 537 PESRAA 542

RESULT 19
US-09-147-052-4
; Sequence 4, Application US/09147052
; Patent No. US20010014335A1
; GENERAL INFORMATION:
; APPLICANT: SAITOH, Shuji
; APPLICANT: TSUZUKI, Yoshiaki
; APPLICANT: YAMAGIDA, No. US20010014335A1oru
; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,
; FILE REFERENCE: 981167
; CURRENT APPLICATION NUMBER: US/09/147,052
; CURRENT FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: JP 08-103548
; PRIOR FILING DATE: 1996-03-29
; PRIOR APPLICATION NUMBER: PCT/JP97/01084
; PRIOR FILING DATE: 1997-03-28
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 4
; LENGTH: 1086
; TYPE: PRT
; ORGANISM: hybrid
US-09-147-052-4

Query Match
40.0%; Score 6; DB 10; Length 1086;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 TPESRA 14
DB 174 TPESRA 179

RESULT 20
US-09-866-510-14
; Sequence 14, Application US/09866510
; Patent No. US2002011304A1
; GENERAL INFORMATION:
; APPLICANT: KAZLAUSKAS, ANDRIUS
; APPLICANT: IKUNO, YASUSHI
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
; FILE REFERENCE: ERM-104.01
; CURRENT APPLICATION NUMBER: US/09/866,510
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/250,747
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/289,103
```

```

; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 14
; LENGTH: 1090
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-866-510-14

Query Match
40.0%; Score 6; DB 10; Length 1090;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPH 6
DB 656 SHLGPH 661

RESULT 21
US-09-955-363-2
; Sequence 2, Application US/09955363
; Patent No. US20020173621A1
; GENERAL INFORMATION:
; APPLICANT: Siedziwski Ph.D., Andrzej Z
; APPLICANT: Bell, Lillian A.
; APPLICANT: Kindsvogel Ph.D., Wayne R.
; TITLE OF INVENTION: METHODS OF PRODUCING SECRETED RECEPTOR ANALOGS
; AND BIOLOGICALLY ACTIVE DIMERIZED POLYPEPTIDE
; FUSIONS
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESS: Seed and Berry
; STREET: 6300 Columbia Center
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/955,363
; FILING DATE: 18-Sep-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/634,510
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/347,291
; FILING DATE: 02-MAY-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, J.D., David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 990008.446C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-622-4900
; TELEFAX: 206-682-6031
; TELEX: 3723836
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1106 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-955-363-2

Query Match
40.0%; Score 6; DB 9; Length 1106;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPH 6
```

DB 656 SHLGPB 661

RESULT 22

US-09-866-510-16
; Sequence 16, Application US/09866510
; Patent No. US20020111304A1
; GENERAL INFORMATION:
; APPLICANT: KAZLAUSKAS, ANDRIUS
; APPLICANT: IKUNO, YASUSHI
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
; FILE REFERENCE: ERM-104.01
; CURRENT APPLICATION NUMBER: US/09/866,510
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/250,747
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/289,103
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 1106
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-866-510-16

Query Match

Best Local Similarity 40.0%; Score 6; DB 10; Length 1106;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPB 6

DB 656 SHLGPB 661

RESULT 23

US-09-866-510-18
; Sequence 18, Application US/09866510
; Patent No. US20020111304A1
; GENERAL INFORMATION:
; APPLICANT: KAZLAUSKAS, ANDRIUS
; APPLICANT: IKUNO, YASUSHI
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
; FILE REFERENCE: ERM-104.01
; CURRENT APPLICATION NUMBER: US/09/866,510
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/250,747
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/289,103
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 1106
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-866-510-18

Query Match

Best Local Similarity 40.0%; Score 6; DB 10; Length 1106;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPB 6

DB 656 SHLGPB 661

RESULT 24

US-09-866-510-20
; Sequence 20, Application US/09866510
; Patent No. US20020111304A1
; GENERAL INFORMATION:

; APPLICANT: KAZLAUSKAS, ANDRIUS

; APPLICANT: IKUNO, YASUSHI

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES

; FILE REFERENCE: ERM-104.01

; CURRENT APPLICATION NUMBER: US/09/866,510

; CURRENT FILING DATE: 2001-05-25

; PRIOR APPLICATION NUMBER: 60/250,747

; PRIOR FILING DATE: 2000-12-01

; PRIOR APPLICATION NUMBER: 60/289,103

; PRIOR FILING DATE: 2001-05-07

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 20

; LENGTH: 1106

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-866-510-20

QY 1 SHLGPB 6

DB 656 SHLGPB 661

RESULT 25

US-09-866-510-22
; Sequence 22, Application US/09866510
; Patent No. US20020111304A1
; GENERAL INFORMATION:
; APPLICANT: KAZLAUSKAS, ANDRIUS
; APPLICANT: IKUNO, YASUSHI
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
; FILE REFERENCE: ERM-104.01
; CURRENT APPLICATION NUMBER: US/09/866,510
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/250,747
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/289,103
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 1106
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-866-510-22

Query Match

Best Local Similarity 40.0%; Score 6; DB 10; Length 1106;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPB 6

DB 656 SHLGPB 661

RESULT 26

US-09-734-520-75
; Sequence 75, Application US/09734520
; Patent No. US20020115173A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: SHORT PEPTIDES FROM THE A-REGION OF
; TITLE OF INVENTION: PROTEIN KINASES WHICH SELECTIVELY MODULATE PROTEIN KINASE
; FILE REFERENCE: 1242.2003-000
; CURRENT APPLICATION NUMBER: US/09/734,520
; CURRENT FILING DATE: 2000-12-11
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSeq for Windows Version 4.0

```
; SEQ ID NO 75
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PDGFR-b
; NAME/KEY: MYRISTATE
; LOCATION: (1)...(0)
; NAME/KEY: AMIDATION
; LOCATION: (0)...(12)
US-09-734-520-75
```

```
Query Match          33.3%; Score 5; DB 10; Length 12;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      2 HLGPH 6
        |||||
DB       2 HLGPH 6
```

```
RESULT 27
US-09-734-520-76
; Sequence 76, Application US/09734520
; Patent No. US2002015173A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: SHORT PEPTIDES FROM THE A-REGION OF
; TITLE OF INVENTION: PROTEIN KINASES WHICH SELECTIVELY MODULATE PROTEIN KINASE
; FILE REFERENCE: 1242.2003-000
; CURRENT APPLICATION NUMBER: US/09/734,520
; CURRENT FILING DATE: 2000-12-11
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PDGFR-a
; NAME/KEY: MYRISTATE
; LOCATION: (1)...(0)
; NAME/KEY: AMIDATION
; LOCATION: (0)...(12)
US-09-734-520-76
```

```
Query Match          33.3%; Score 5; DB 10; Length 12;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      2 HLGPH 6
        |||||
DB       2 HLGPH 6
```

```
RESULT 28
US-10-012-034A-75
; Sequence 75, Application US/10012034A
; Patent No. US20020137141A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: SHORT PEPTIDES FROM THE A-REGION OF
; TITLE OF INVENTION: PROTEIN KINASES WHICH SELECTIVELY MODULATE PROTEIN KINASE
; FILE REFERENCE: BEN-SASSON-5A
; CURRENT APPLICATION NUMBER: US/10/012,034A
; CURRENT FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 09/734,520
; PRIOR FILING DATE: 2000-12-11
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 75
```

```
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PDGFR-b
; NAME/KEY: MYRISTATE
; LOCATION: (1)...(0)
; FEATURE:
; NAME/KEY: AMIDATION
; LOCATION: (0)...(12)
US-10-012-034A-75
```

```
Query Match          33.3%; Score 5; DB 12; Length 12;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      2 HLGPH 6
        |||||
DB       2 HLGPH 6
```

```
RESULT 29
US-10-012-034A-76
; Sequence 76, Application US/10012034A
; Patent No. US20020137141A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: SHORT PEPTIDES FROM THE A-REGION OF
; TITLE OF INVENTION: PROTEIN KINASES WHICH SELECTIVELY MODULATE PROTEIN KINASE
; FILE REFERENCE: BEN-SASSON-5A
; CURRENT APPLICATION NUMBER: US/10/012,034A
; CURRENT FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 09/734,520
; PRIOR FILING DATE: 2000-12-11
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PDGFR-a
; NAME/KEY: MYRISTATE
; LOCATION: (1)...(0)
; FEATURE:
; NAME/KEY: AMIDATION
; LOCATION: (0)...(12)
US-10-012-034A-76
```

```
Query Match          33.3%; Score 5; DB 12; Length 12;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      2 HLGPH 6
        |||||
DB       2 HLGPH 6
```

```
RESULT 30
US-09-734-520-15
; Sequence 15, Application US/09734520
; Patent No. US2002015173A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: SHORT PEPTIDES FROM THE A-REGION OF
; TITLE OF INVENTION: PROTEIN KINASES WHICH SELECTIVELY MODULATE PROTEIN KINASE
; FILE REFERENCE: 1242.2003-000
; CURRENT APPLICATION NUMBER: US/09/734,520
; CURRENT FILING DATE: 2000-12-11
```

```
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 15
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: PDGFR-a
US-09-734-520-15

Query Match          33.3%; Score 5; DB 10; Length 19;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HLGP 6
DB 6 HLGP 10

RESULT 31
US-10-012-034A-15
; Sequence 15, Application US/10012034A
; Patent No. US20020137141A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: SHORT PEPTIDES FROM THE A-REGION OF
; TITLE OF INVENTION: PROTEIN KINASES WHICH SELECTIVELY MODULATE PROTEIN KINASE
; FILE REFERENCE: ACTIVITY
; CURRENT APPLICATION NUMBER: US/10/012,034A
; CURRENT FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 09/734,520
; PRIOR FILING DATE: 2000-12-11
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 15
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PDGFR-a
; NAME/KEY: MYRISTATE
; LOCATION: (1)...(0)
; FEATURE:
; NAME/KEY: AMIDATION
; LOCATION: (0)...(19)
US-10-012-034A-15

Query Match          33.3%; Score 5; DB 12; Length 19;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HLGP 6
DB 6 HLGP 10

RESULT 32
US-09-864-761-44095
; Sequence 44095, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecmlca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
```

```
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO: 44095
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC003086.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.76
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.76
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.75
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.68
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.73
US-09-864-761-44095

Query Match          33.3%; Score 5; DB 10; Length 34;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTP 11
DB 23 RSTP 27

RESULT 33
US-09-774-639-274
; Sequence 274, Application US/09774639
; Publication No. US20030003555A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 90 Human Secreted Proteins
; FILE REFERENCE: P2013P1
; CURRENT APPLICATION NUMBER: US/09/774,639
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/244,112
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 371
```

```

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 274
; LENGTH: 39
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-774-639-274

Query Match
Best Local Similarity 100.0%; Score 5; DB 9; Length 39;
Pred. No. 48;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 HRSTP 10
DB 10 HRSTP 14

RESULT 34
US-09-774-639-351
; Sequence 351, Application US/09774639
; Publication No. US20030003555a1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 90 Human Secreted Proteins
; FILE REFERENCE: P2013p1
; CURRENT APPLICATION NUMBER: US/09/774,639
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/244,112
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 371
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 351
; LENGTH: 39
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-774-639-351

Query Match
Best Local Similarity 100.0%; Score 5; DB 9; Length 39;
Pred. No. 48;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 HRSTP 10
DB 10 HRSTP 14

RESULT 35
US-09-864-761-44670
; Sequence 44670, Application US/09864761
; Patent No. US20020048763a1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
```

```

; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 44670
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC021383.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.66
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.72
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.67
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.76
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.64
; OTHER INFORMATION: EST_HUMAN HIT: BE256045.1, EVALUO 3.00e+00
US-09-864-761-44670

Query Match
Best Local Similarity 100.0%; Score 5; DB 10; Length 48;
Pred. No. 57;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GPHRS 8
DB 20 GPHRS 24

RESULT 36
US-09-764-847-695
; Sequence 695, Application US/09764847
; Patent No. US20020132767a1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 695
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (59)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-847-695

Query Match
Best Local Similarity 100.0%; Score 5; DB 10; Length 61;
Pred. No. 71;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Oy 1 SHUGP 5
|||||
Db 22 SHUGP 26

RESULT 37
US-09-864-761-33764
; Sequence 33764, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33764
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL035610.2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 5.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.9
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.8
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 5.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 6.8
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 9.3

; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 7
; OTHER INFORMATION: SWISSPROT HIT: P43527, EVALU6 6.10e+00
; OTHER INFORMATION: EST_HUMAN HIT: AA746130.1, EVALU6 1.00e-50
US-09-864-761-33764

Query Match 33.3%; Score 5; DB 10; Length 91;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 4 GPHRS 8
|||||
Db 10 GPHRS 14

RESULT 38
US-09-738-626-5971
; Sequence 5971, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAMA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 5971
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5971

Query Match 33.3%; Score 5; DB 9; Length 97;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 9 TPESR 13
|||||
Db 42 TPESR 46

RESULT 39
US-09-852-137-2
; Sequence 2, Application US/09852137
; Patent No. US2002005154A1
; GENERAL INFORMATION:
; APPLICANT: MCKEL, BETTINA
; APPLICANT: WEISSENBOHN, ANKE
; APPLICANT: PEEFERLE, WALTER
; APPLICANT: SAHM, HERMANN
; APPLICANT: EGGLING, LOTHAR
; APPLICANT: NAMPOOTHIRI, MADHAVAN
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE ACP GENE
; FILE REFERENCE: 21123/280306/MAS
; CURRENT APPLICATION NUMBER: US/09/852,137
; CURRENT FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: DE 100 23 400.3
; PRIOR FILING DATE: 2000-12-05

```
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-852-137-2
```

```
Query Match          33.3%; Score 5; DB 10; Length 97;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      9 TPESR 13
        |||||
DB      42 TPESR 46
```

```
RESULT 40
US-09-864-761-41798
; Sequence 41798, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 41798
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```
; FEATURE:
; OTHER INFORMATION: MAP TO AC020557.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.8
; OTHER INFORMATION: EST HUMAN HIT: D54392.1, EVALUATE 4.60e+00
; OTHER INFORMATION: SWISSPROT HIT: Q98910, EVALUATE 1.10e+00
US-09-864-761-41798
```

```
Query Match          33.3%; Score 5; DB 10; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      11 ESRAA 15
        |||||
DB      82 ESRAA 86
```

```
RESULT 41
US-09-736-457-1671
; Sequence 1671, Application US/09736457
; Patent No. US20020168637A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darriick
; APPLICANT: Retteer, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aljun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1671
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-736-457-1671

Query Match          33.3%; Score 5; DB 9; Length 109;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      3 LGPHR 7
        |||||
DB      54 LGPHR 58
```

```
RESULT 42
US-09-902-941-1671
; Sequence 1671, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retteer, Marc W.
; APPLICANT: Mamerakis, Margarita
; APPLICANT: Carter, Darriick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
```



```
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1671
; LENGTH: 109
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-902-941-1671
```

```
Query Match          33.3%; Score 5; DB 9; Length 109;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 3 LGPHR 7
    11111
DB 54 LGPHR 58
```

```
RESULT 43
US-09-849-626-1671
; Sequence 1671, Application US/09849626
; Publication No. US20020197669A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya
; APPLICANT: Fanger, Gary
; APPLICANT: Wang, Aijun
; APPLICANT: Wang, Tonglong
; APPLICANT: Switzer, Anne
; APPLICANT: McNeill, Patricia
; APPLICANT: Clapper, Jonathan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C16
; CURRENT APPLICATION NUMBER: US/09/849,626
; CURRENT FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 1926
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 1671
; LENGTH: 109
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-849-626-1671
```

```
Query Match          33.3%; Score 5; DB 9; Length 109;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 3 LGPHR 7
    11111
DB 54 LGPHR 58
```

```
RESULT 44
US-09-925-300-1175
; Sequence 1175, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1175
```

```
; LENGTH: 114
; TYPE: PRF
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (50)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1175
```

```
Query Match          33.3%; Score 5; DB 10; Length 114;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 11 ESRRA 15
    11111
DB 93 ESRRA 97
```

```
RESULT 45
US-10-004-551-12
; Sequence 12, Application US/10004551
; Publication No. US20030004310A1
; GENERAL INFORMATION:
; APPLICANT: SHINKERS, RICHARD A
; APPLICANT: FERNANDES, ELMA
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ENCODED THEREBY
; FILE REFERENCE: 15966-359
; CURRENT APPLICATION NUMBER: US/10/004,551
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 09/635,949
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 121
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-004-551-12
```

```
Query Match          33.3%; Score 5; DB 9; Length 121;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 7 RSTPE 11
    11111
DB 55 RSTPE 59
```

```
RESULT 46
US-09-738-626-6171
; Sequence 6171, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: YOKOI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHITO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
```

```
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6171
; LENGTH: 126
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6171
```

```
Query Match          33.3%; Score 5; DB 9; Length 126;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 5 PHRST 9
    |||||
Db 104 PHRST 108
```

```
RESULT 47
US-09-734-569-86
; Sequence 86, Application US/09734569
; Patent No. US20020064816A1
; GENERAL INFORMATION:
; APPLICANT: Lerchl, Jens
; APPLICANT: Renz, Andreas
; APPLICANT: Ehrhardt, Thomas
; APPLICANT: Reinold, Andreas
; APPLICANT: Cirpus, Petra
; APPLICANT: Bischoff, Friedrich
; APPLICANT: Frank, Markus
; APPLICANT: Freund, Annette
; APPLICANT: Duwenig, Elke
; APPLICANT: Schmidt, Ralf-Michael
; APPLICANT: Reski, Ralf
; TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins involved
; FILE REFERENCE: BASF-NAE-1332-99-US
; CURRENT APPLICATION NUMBER: US/09/734,569
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 60/171,101
; PRIOR FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: PatentIn Ver. 2.1/WordPerfect 6.1
; SEQ ID NO 86
; LENGTH: 126
; TYPE: PRT
; ORGANISM: Physcomitrella patens
US-09-734-569-86
```

```
Query Match          33.3%; Score 5; DB 10; Length 126;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 11 ESRAA 15
    |||||
Db 86 ESRAA 90
```

```
RESULT 48
US-09-764-868-781
; Sequence 781, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT32
; CURRENT APPLICATION NUMBER: US/09/764,868
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 781
; LENGTH: 129
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (89)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (97)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (105)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (112)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-868-781
```

```
Query Match          33.3%; Score 5; DB 9; Length 129;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 4 GPHRS 8
    |||||
Db 100 GPHRS 104
```

```
RESULT 49
US-09-738-626-6788
; Sequence 6788, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIRO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6788
; LENGTH: 130
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6788
```

```
Query Match          33.3%; Score 5; DB 9; Length 130;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 7 RSTPE 11
    |||||
Db 17 RSTPE 21
```

```
RESULT 50
US-09-764-864-1425
; Sequence 1425, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
```

```

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1425
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (63)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (114)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (126)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (127)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1425

Query Match
Best Local Similarity 33.3%; Score 5; DB 10; Length 134;
Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STEPS 12
|11111
DB 81 STEPS 85

RESULT 51
US-09-764-868-1017
; Sequence 1017, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1017
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (134)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (140)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (142)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-868-1017

Query Match
Best Local Similarity 33.3%; Score 5; DB 9; Length 146;
Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STEPS 12
|11111
DB 62 STEPS 66
```

```

RESULT 52
US-09-738-626-4628
; Sequence 4628, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; Prior application data removed - refer to PALM or file wrapper
; Prior application number: JP 99/377484
; Prior filing date: 1999-12-16
; Prior application number: JP 00/159162
; Prior filing date: 2000-04-07
; Prior application number: JP 00/280988
; Prior filing date: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patentln Ver. 3.0
; SEQ ID NO 4628
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4628

Query Match
Best Local Similarity 33.3%; Score 5; DB 9; Length 149;
Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 HRSTP 10
|11111
DB 29 HRSTP 33

RESULT 53
US-09-747-155-243
; Sequence 243, Application US/09747155
; Patent No. US20020151692A1
; GENERAL INFORMATION:
; APPLICANT: Rouquier, Sylvie
; APPLICANT: Giorzi, Dominique
; TITLE OF INVENTION: No. US20020151692A1 Polypeptides and Nucleic Acids Encoding
; FILE REFERENCE: 19904-008 (C009B6834US)
; CURRENT APPLICATION NUMBER: US/09/747,155
; CURRENT FILING DATE: 2000-12-21
; Prior application number: 60/171,746
; Prior filing date: 1999-12-22
; NUMBER OF SEQ ID NOS: 431
; SOFTWARE: Patentln version 3.0
; SEQ ID NO 243
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Eulemur fulvus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(485)
; OTHER INFORMATION: Taxon = 13515; gene = EFU145; Accession DDBJ/EMBL/GenBank = AF
US-09-747-155-243

Query Match
Best Local Similarity 33.3%; Score 5; DB 10; Length 161;
Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRAA 15
|11111
```

Db 148 ESRAA 152

RESULT 54
US-09-736-457-798

```
; Sequence 798, Application US/09736457
; Patent No. US20020168637A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aljun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-736-457-798
```

```
Query Match          33.3%; Score 5; DB 9; Length 164;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 3 LGPFR 7
Db 109 LGPFR 113

RESULT 55

```
; Sequence 798, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 798
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-902-941-798
```

```
Query Match          33.3%; Score 5; DB 9; Length 164;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 3 LGPFR 7
Db 109 LGPFR 113

RESULT 56

```
US-09-849-626-798
; Sequence 798, Application US/09849626
; Publication No. US20020197669A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya
; APPLICANT: Fanger, Gary
; APPLICANT: Wang, Aljun
; APPLICANT: Wang, Tonglong
; APPLICANT: Switzer, Anne
; APPLICANT: McNeill, Patricia
; APPLICANT: Clapper, Jonathan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C16
; CURRENT APPLICATION NUMBER: US/09/849,626
; CURRENT FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 1926
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 798
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-849-626-798
```

```
Query Match          33.3%; Score 5; DB 9; Length 164;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 3 LGPFR 7
Db 109 LGPFR 113

RESULT 57

```
; Sequence 49050, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
```

```

; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 49050
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO 282214.22
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.6
; OTHER INFORMATION: EST_HUMAN HIT: A1146468.1, EVALU6 4.00e-03
US-09-864-761-49050
```

```
Query Match          33.3%; Score 5; DB 10; Length 168;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      2 LGPBR 6
        |||||
DB      94 HLGPH 98
```

```
RESULT 58
US-09-736-457-1678
; Sequence 1678, Application US/09736457
; Patent No. US20020168637A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodges, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aljun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1678
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-736-457-1678
```

```
Query Match          33.3%; Score 5; DB 9; Length 177;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      3 LGPBR 7
        |||||
DB      122 LGPBR 126
```

```
RESULT 59
US-09-902-941-1678
; Sequence 1678, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Matanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1678
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-902-941-1678
```

```
Query Match          33.3%; Score 5; DB 9; Length 177;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      3 LGPBR 7
        |||||
DB      122 LGPBR 126
```

```
RESULT 60
US-09-849-626-1678
; Sequence 1678, Application US/09849626
; Publication No. US20020197669A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya
; APPLICANT: Fanger, Gary
; APPLICANT: Wang, Aljun
; APPLICANT: Wang, Tonglong
; APPLICANT: Switzer, Anne
; APPLICANT: McNeill, Patricia
; APPLICANT: Clapper, Jonathan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C16
; CURRENT APPLICATION NUMBER: US/09/849,626
; CURRENT FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 1926
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1678
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-849-626-1678
```

```
Query Match          33.3%; Score 5; DB 9; Length 177;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      3 LGPBR 7
        |||||
DB      122 LGPBR 126
```

```
RESULT 61
US-10-093-766-52
; Sequence 52, Application US/10093766
; Publication No. US20030013099A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; APPLICANT: Jones, David A.
; APPLICANT: Karpf, Adam R.
; TITLE OF INVENTION: GENES REGULATED BY DNA METHYLATION IN COLON TUMORS
; FILE REFERENCE: PA-0047 US
; CURRENT APPLICATION NUMBER: US/10/093,766
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PERL Program
; SEQ ID NO 52
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No. US20030013099A1 2238363CD1
US-10-093-766-52

Query Match          33.3%; Score 5; DB 9; Length 181;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPFR 7
Db 17 LGPFR 21

RESULT 62
US-09-811-284-176
; Sequence 176, Application US/09811284
; Patent No. US20020058306A1
; GENERAL INFORMATION:
; APPLICANT: Vogel, Gabriel
; TITLE OF INVENTION: No. US20020058306A1el G Protein-Coupled Receptors
; FILE REFERENCE: 00167US1
; CURRENT APPLICATION NUMBER: US/09/811,284
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/189,783
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/189,907
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/189,918
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/189,960
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/189,917
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/192,945
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/192,916
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/192,923
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/192,933
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/192,830
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/192,234
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/192,155
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,935
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 258
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 176
; LENGTH: 186
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
US-09-811-284-176

Query Match          33.3%; Score 5; DB 10; Length 186;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRA 14
Db 111 PESRA 115

RESULT 63
US-09-996-015-43
; Sequence 43, Application US/09996015
; Publication No. US20030032166A1
; GENERAL INFORMATION:
; APPLICANT: Quinn, Kerry E.
; APPLICANT: Pena, Carol A. E.
; APPLICANT: Li, Li
; APPLICANT: Spaderna, Steven K.
; TITLE OF INVENTION: Aortic Carboxypeptidase-Like Proteins and Nucleic Acids
; FILE REFERENCE: 15966-581 CIP
; CURRENT APPLICATION NUMBER: US/09/996,015
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 09/641,741
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 60/159,613
; PRIOR FILING DATE: 1999-10-14
; PRIOR APPLICATION NUMBER: 60/175,534
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 60/224,086
; PRIOR FILING DATE: 2000-08-09
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-996-015-43

Query Match          33.3%; Score 5; DB 9; Length 192;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPFR 7
Db 140 LGPFR 144

RESULT 64
US-09-996-015-44
; Sequence 44, Application US/09996015
; Publication No. US20030032166A1
; GENERAL INFORMATION:
; APPLICANT: Quinn, Kerry E.
; APPLICANT: Pena, Carol A. E.
; APPLICANT: Li, Li
; APPLICANT: Spaderna, Steven K.
; TITLE OF INVENTION: Aortic Carboxypeptidase-Like Proteins and Nucleic Acids
; FILE REFERENCE: 15966-581 CIP
; CURRENT APPLICATION NUMBER: US/09/996,015
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 09/641,741
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 60/159,613
; PRIOR FILING DATE: 1999-10-14
; PRIOR APPLICATION NUMBER: 60/175,534
; PRIOR FILING DATE: 2000-01-11
```

```
; PRIOR APPLICATION NUMBER: 60/224,086
; PRIOR FILING DATE: 2000-08-09
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-996-015-44
```

```
Query Match          33.3%; Score 5; DB 9; Length 193;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      3 LGPFR 7
        |||||
Db      140 LGPFR 144
```

```
RESULT 65
US-09-996-015-47
; Sequence 47, Application US/09996015
; Publication No. US2003003216A1
; GENERAL INFORMATION:
; APPLICANT: Quinn, Kerry E.
; APPLICANT: Pena, Carol A. E.
; APPLICANT: Li, Li
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Leite, Mario W.
; TITLE OF INVENTION: Aortic Carboxypeptidase-Like Proteins and Nucleic Acids
; FILE REFERENCE: 15966-581 CIP
; CURRENT APPLICATION NUMBER: US/09/996,015
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 09/641,741
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 60/159,613
; PRIOR FILING DATE: 1999-10-14
; PRIOR APPLICATION NUMBER: 60/175,534
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 60/224,086
; PRIOR FILING DATE: 2000-08-09
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 47
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-996-015-47
```

```
Query Match          33.3%; Score 5; DB 9; Length 193;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      3 LGPFR 7
        |||||
Db      140 LGPFR 144
```

```
RESULT 66
US-09-996-015-8
; Sequence 8, Application US/09996015
; Publication No. US2003003216A1
; GENERAL INFORMATION:
; APPLICANT: Quinn, Kerry E.
; APPLICANT: Pena, Carol A. E.
; APPLICANT: Li, Li
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Leite, Mario W.
; TITLE OF INVENTION: Aortic Carboxypeptidase-Like Proteins and Nucleic Acids
; FILE REFERENCE: 15966-581 CIP
; CURRENT APPLICATION NUMBER: US/09/996,015
```

```
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 09/641,741
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 60/159,613
; PRIOR FILING DATE: 1999-10-14
; PRIOR APPLICATION NUMBER: 60/175,534
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 60/224,086
; PRIOR FILING DATE: 2000-08-09
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 202
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-996-015-8
```

```
Query Match          33.3%; Score 5; DB 9; Length 202;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      3 LGPFR 7
        |||||
Db      140 LGPFR 144
```

```
RESULT 67
US-09-738-626-5128
; Sequence 5128, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SETKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, MAKOTO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 5128
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5128
```

```
Query Match          33.3%; Score 5; DB 9; Length 205;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      8 STPES 12
        |||||
Db      114 STPES 118
```

```
RESULT 68
US-09-925-301-898
; Sequence 898, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: P106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 898
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (205)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-301-898
```

```

Query Match          33.3%; Score 5; DB 10; Length 232;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      10 PESRA 14
      11111
DB      194 PESRA 198
```

```

RESULT 69
US-09-815-242-5905
; Sequence 5905, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes In
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5905
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5905
```

```

Query Match          33.3%; Score 5; DB 10; Length 248;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      11 ESRA 15
      11111
DB      20 ESRA 24
```

```

RESULT 70
US-09-764-853-658
; Sequence 658, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P1206
; CURRENT APPLICATION NUMBER: US/09/764,853
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 658
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-853-658
```

```

Query Match          33.3%; Score 5; DB 10; Length 269;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      10 PESRA 14
      11111
DB      195 PESRA 199
```

```

RESULT 71
US-09-908-193-40
; Sequence 40, Application US/09908193
; Publication No. US20020192748A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; APPLICANT: SHIMKETS, RICHARD A.
; APPLICANT: ZERHUSEN, BRYAN
; APPLICANT: MALYANKAR, URIEL M.
; APPLICANT: PADIGARU, MURALIDHARA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES ENCODED THEREBY
; FILE REFERENCE: 21402-062
; CURRENT APPLICATION NUMBER: US/09/908,193
; CURRENT FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: 60/220,273
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: 60/221,650
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/221,233
; PRIOR FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: 60/220,912
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/218,875
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/218,870
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/218,901
; PRIOR FILING DATE: 2000-07-18
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-908-193-40
```

```

Query Match          33.3%; Score 5; DB 9; Length 271;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```


QY 6 HRSTP 10
|||||
Db 79 HRSTP 83

RESULT 72

US-09-738-626-4071
; Sequence 4071, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patentln ver. 3.0
; SEQ ID NO 4071
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4071

Query Match 33.3%; Score 5; DB 9; Length 280;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPE 11
|||||
Db 26 RSTPE 30

RESULT 73

US-09-925-300-1344
; Sequence 1344, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1344
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-300-1344

Query Match 33.3%; Score 5; DB 10; Length 310;
Best Local Similarity 100.0%; Pred. No. 3e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHLGP 5
|||||
Db 88 SHLGP 92

RESULT 74

US-09-764-868-1009
; Sequence 1009, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PPT32
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1009
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-868-1009

Query Match 33.3%; Score 5; DB 9; Length 311;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HLGPH 6
|||||
Db 105 HLGPH 109

RESULT 75

US-09-791-932-63
; Sequence 63, Application US/09791932
; Publication No. US20030003451A1
; GENERAL INFORMATION:
; APPLICANT: Vogel, Gabriel
; APPLICANT: Parodi, Luis A.
; APPLICANT: Hiesch, Ronald R.
; APPLICANT: Lind, Peter
; APPLICANT: Kayles, Paul S.
; APPLICANT: Ruff, Valerie
; APPLICANT: Huff, Rita M.
; TITLE OF INVENTION: No. US20030003451A1el G Protein-Coupled Receptors Cross-Ref
; FILE REFERENCE: 00325 US1
; CURRENT APPLICATION NUMBER: US/09/791,932
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/184,305
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/184,304
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/184,303
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/184,397
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/184,247
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/188,880
; PRIOR FILING DATE: 2000-03-13
; PRIOR APPLICATION NUMBER: 60/217,369
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/217,370
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/218,492
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: 60/186,810
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/188,064

;; PRIOR FILING DATE: 2000-03-09
;; PRIOR APPLICATION NUMBER: 60/186,457
;; PRIOR FILING DATE: 2000-03-02
;; PRIOR APPLICATION NUMBER: 60/213,861
;; PRIOR FILING DATE: 2000-06-23
;; PRIOR APPLICATION NUMBER: 60/194,344
;; PRIOR FILING DATE: 2000-04-03
;; PRIOR APPLICATION NUMBER: 60/218,337
;; PRIOR FILING DATE: 2000-07-14
;; NUMBER OF SEQ ID NOS: 184
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO: 63
;; LENGTH: 313
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-791-932-63

Query Match 33.3%; Score 5; DB 9; Length 313;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 GPHRS 8
|||||
DB 106 GPHRS 110

RESULT 76
US-09-736-457-1863
;; Sequence 1863, Application US/09736457
;; Patent No. US20020168637A1
;; GENERAL INFORMATION:
;; APPLICANT: Wang, Tongtong
;; APPLICANT: Bangur, Chaitanya S.
;; APPLICANT: Lodes, Michael A.
;; APPLICANT: Fanger, Gary
;; APPLICANT: Vedvick, Tom
;; APPLICANT: Carter, Darriek
;; APPLICANT: Retter, Marc
;; APPLICANT: Mannion, Jane
;; APPLICANT: Fan, Aijun
;; APPLICANT: Wang, Aijun
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
;; FILE REFERENCE: 210121.478C15
;; CURRENT APPLICATION NUMBER: US/09/736,457
;; CURRENT FILING DATE: 2000-12-13
;; NUMBER OF SEQ ID NOS: 1864
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO: 1863
;; LENGTH: 314
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-736-457-1863

Query Match 33.3%; Score 5; DB 9; Length 314;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHR 7
|||||
DB 259 LGPHR 263

RESULT 77
US-09-902-941-1863
;; Sequence 1863, Application US/09902941
;; Patent No. US20020172952A1
;; GENERAL INFORMATION:
;; APPLICANT: Henderson, Robert A.
;; APPLICANT: Wang, Tongtong
;; APPLICANT: Watanabe, Yoshihiro
;; APPLICANT: Johnson, Jeffrey C.
;; APPLICANT: Retter, Marc W.

;; APPLICANT: Marnetakis, Margarita
;; APPLICANT: Carter, Darriek
;; APPLICANT: Fanger, Gary R.
;; APPLICANT: Vedvick, Thomas S.
;; APPLICANT: Bangur, Chaitanya S.
;; APPLICANT: McNabb, Andria
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
;; FILE REFERENCE: 210121.478C17
;; CURRENT APPLICATION NUMBER: US/09/902,941
;; CURRENT FILING DATE: 2001-07-10
;; NUMBER OF SEQ ID NOS: 2002
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 1863
;; LENGTH: 314
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-902-941-1863

Query Match 33.3%; Score 5; DB 9; Length 314;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHR 7
|||||
DB 259 LGPHR 263

RESULT 78
US-09-849-626-1863
;; Sequence 1863, Application US/09849626
;; Publication No. US2002019769A1
;; GENERAL INFORMATION:
;; APPLICANT: Bangur, Chaitanya
;; APPLICANT: Fanger, Gary
;; APPLICANT: Wang, Aijun
;; APPLICANT: Wang, Tongtong
;; APPLICANT: Switzer, Anne
;; APPLICANT: McNeill, Patricia
;; APPLICANT: Clapper, Jonathan
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
;; FILE REFERENCE: 210121.478C16
;; CURRENT APPLICATION NUMBER: US/09/849,626
;; CURRENT FILING DATE: 2001-05-03
;; NUMBER OF SEQ ID NOS: 1926
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO: 1863
;; LENGTH: 314
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-849-626-1863

Query Match 33.3%; Score 5; DB 9; Length 314;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHR 7
|||||
DB 259 LGPHR 263

RESULT 79
US-09-975-719-23
;; Sequence 23, Application US/09975719
;; Publication No. US20030022349A1
;; GENERAL INFORMATION:
;; APPLICANT: Ausubel, Frederick M.
;; APPLICANT: Rahme, Laurence G.
;; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
;; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
;; FILE REFERENCE: 00786/361003
;; CURRENT APPLICATION NUMBER: US/09/975,719

; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 09/199,637
; PRIOR FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: US 60/066,517
; PRIOR FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-975-719-23

Query Match 33.3%; Score 5; DB 9; Length 318;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 HRSTP 10
Db 13 HRSTP 17

RESULT 80
US-09-908-193-39
; Sequence 39, Application US/09908193
; Publication No. US20020192748A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; APPLICANT: SHIMETS, RICHARD A.
; APPLICANT: ZERHUSEN, BRYAN
; APPLICANT: MALYANKAR, URIEL M.
; APPLICANT: PADIGARU, MURALIDHARA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES ENCODED THEREBY
; FILE REFERENCE: 21402-062
; CURRENT APPLICATION NUMBER: US/09/908,193
; CURRENT FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: 60/220,273
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: 60/221,650
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/221,233
; PRIOR FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: 60/220,912
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/218,875
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/218,870
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/218,901
; PRIOR FILING DATE: 2000-07-18
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Mus musculus
; NAME/KEY: MOD_RESS
; LOCATION: (314)
; OTHER INFORMATION: Any Amino Acid
US-09-908-193-39

Query Match 33.3%; Score 5; DB 9; Length 327;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 HRSTP 10
Db 136 HRSTP 140

RESULT 81
US-09-738-626-4350

; Sequence 4350, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOI, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4350
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4350

Query Match 33.3%; Score 5; DB 9; Length 329;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGP 5
Db 288 SHLGP 292

RESULT 82
US-09-886-055-211
; Sequence 211, Application US/09886055
; Patent No. US20020132273A1
; GENERAL INFORMATION:
; APPLICANT: STRYER, LUBERT
; APPLICANT: ZOZULYA, SERGEY
; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
; TITLE OF INVENTION: BIOSENSORS OF CHEMICAL STIMULANTS
; FILE REFERENCE: 078003-0277150
; CURRENT APPLICATION NUMBER: US/09/886,055
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,812
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 522
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 211
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-886-055-211

Query Match 33.3%; Score 5; DB 10; Length 330;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRAA 15
Db 288 ESRAA 292

RESULT 83
US-09-908-193-37

```
; Sequence 37, Application US/09908193
; Publication No. US20020192748A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; APPLICANT: SHIMKETS, RICHARD A.
; APPLICANT: ZERHUSEN, BRYAN
; APPLICANT: MALYANKAR, URIEL M.
; APPLICANT: PADIGARU, MURALIDHARA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES ENCODED THEREBY
; FILE REFERENCE: 21402-062
; CURRENT APPLICATION NUMBER: US/09/908,193
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: 60/220,273
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: 60/221,650
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/221,233
; PRIOR FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: 60/220,912
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/220,912
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/218,875
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/218,870
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/218,901
; PRIOR FILING DATE: 2000-07-18
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-908-193-37

Query Match          33.3%; Score 5; DB 9; Length 335;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      6 HRSTP 10
      |||||
Db      143 HRSTP 147

RESULT 84
US-09-908-193-38
; Sequence 38, Application US/09908193
; Publication No. US20020192748A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; APPLICANT: SHIMKETS, RICHARD A.
; APPLICANT: ZERHUSEN, BRYAN
; APPLICANT: MALYANKAR, URIEL M.
; APPLICANT: PADIGARU, MURALIDHARA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES ENCODED THEREBY
; FILE REFERENCE: 21402-062
; CURRENT APPLICATION NUMBER: US/09/908,193
; CURRENT FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: 60/220,273
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: 60/221,650
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/221,233
; PRIOR FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: 60/220,912
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/218,875
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/218,870
; PRIOR APPLICATION NUMBER: 60/218,901
; PRIOR FILING DATE: 2000-07-18
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 38
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-908-193-38

Query Match          33.3%; Score 5; DB 9; Length 336;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      6 HRSTP 10
      |||||
Db      144 HRSTP 148

RESULT 85
US-09-815-242-12043
; Sequence 12043, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes In
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12043
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-12043

Query Match          33.3%; Score 5; DB 10; Length 338;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      11 ESRRA 15
      |||||
Db      269 ESRRA 273

RESULT 86
US-10-114-170-263
; Sequence 263, Application US/10114170
; Publication No. US20030023075A1
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; APPLICANT: Burland, Valerie
; APPLICANT: Perna, Nicole T.
; APPLICANT: Plunkett, Guy
```

```

;                               Welch, Red
; TITLE OF INVENTION: No. US20030023075a1el Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Plinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 Inch. 1.44mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/114,170
; FILING DATE: 01-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 09/453,702
; FILING DATE: 03-DEC-1999
; APPLICATION NUMBER: 60/110,955
; FILING DATE: 04-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 263:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 348 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 263:
US-10-114-170-263

Query Match          33.3%; Score 5; DB 9; Length 348;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRAA 15
Db 160 ESRAA 164

RESULT 87
US-09-908-805B-25
; Sequence 25, Application US/099080805B
; Patent No. US20020147307A1
; GENERAL INFORMATION:
; APPLICANT: HILTON, Douglas J
; APPLICANT: ALEXANDER, Warren S
; APPLICANT: VINEY, Elizabeth M
; APPLICANT: WILSON, Tracey A
; APPLICANT: RICHARDSON, Rachael T
; APPLICANT: STARR, Robyn
; APPLICANT: NICHOLSON, Sandra E
; APPLICANT: METCALF, Donald
; APPLICANT: NICOLA, Nicos A
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC AGENTS
; FILE REFERENCE: 109762
; CURRENT APPLICATION NUMBER: US/09/908,805B
; CURRENT FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 08/962,560
; PRIOR FILING DATE: 1997-10-31
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
```

```

; LENGTH: 350
; TYPE: PRT
; ORGANISM: Mouse
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (167)
; OTHER INFORMATION: Xaa 1s unsure
US-09-908-805B-25

Query Match          33.3%; Score 5; DB 10; Length 350;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 HRSTP 10
Db 158 HRSTP 162

RESULT 88
US-10-108-714-4
; Sequence 4, Application US/10108714
; Patent No. US20020128445A1
; GENERAL INFORMATION:
; APPLICANT: Regan, John W.
; APPLICANT: Gil, Daniel W.
; APPLICANT: Woodward, David F.
; TITLE OF INVENTION: No. US20020128445a1el Human Prostaglandin EP Receptor
; FILE REFERENCE: 17023 DIV CIP
; CURRENT APPLICATION NUMBER: US/10/108,714
; CURRENT FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/267,423
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/239,431
; PRIOR FILING DATE: EARLIER FILING DATE: 1994-05-05
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-714-4

Query Match          33.3%; Score 5; DB 12; Length 358;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRA 14
Db 104 PESRA 108

RESULT 89
US-09-738-626-3763
; Sequence 3763, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SETKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIRO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
```

PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: Patent ver. 3.0
SEQ ID NO 3763
LENGTH: 360
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-3763

Query Match
Best Local Similarity 33.3%; Score 5; DB 9; Length 360;
100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 TPESR 13
|||||
DB 70 TPESR 74

RESULT 90
US-09-870-759-73
Sequence 73, Application US/09870759
Patent No. US20020177551A1
GENERAL INFORMATION:
APPLICANT: TERMAN, David S
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
FILE REFERENCE: 870759
CURRENT APPLICATION NUMBER: US/09/870,759
CURRENT FILING DATE: 2002-01-14
PRIOR APPLICATION NUMBER: US 60/208,128
PRIOR FILING DATE: 2000-05-30
NUMBER OF SEQ ID NOS: 166
SOFTWARE: Patent version 3.1
SEQ ID NO 73
LENGTH: 365
TYPE: PRT
ORGANISM: Mus musculus
US-09-870-759-73

Query Match
Best Local Similarity 33.3%; Score 5; DB 9; Length 365;
100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STPES 12
|||||
DB 192 STPES 196

RESULT 91
US-09-165-522-16
Sequence 16, Application US/09165522
Publication No. US20030023990A1
GENERAL INFORMATION:
APPLICANT: Davis, Roger J.
Flavell, Richard A.
Rakic, Pasko
Whitmarsh, Alan
Kuan, Chia-Yi
Yang, Di
TITLE OF INVENTION: JNK3 MODULATORS AND METHODS OF USE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESSES:
ADDRESS: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95

SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/165,522
FILING DATE: 02-Oct-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/060,995
FILING DATE: 03-Oct-1997
ATTORNEY/AGENT INFORMATION:
NAME: Fasse, J. Peter
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 10363/005001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-09-165-522-16

Query Match
Best Local Similarity 33.3%; Score 5; DB 9; Length 380;
100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STPES 12
|||||
DB 231 STPES 235

RESULT 92
US-09-939-833-6
Sequence 6, Application US/09939833
Patent No. US2002016461A1
GENERAL INFORMATION:
APPLICANT: MOTTIGUE, MICHELE A.
APPLICANT: WICKERSHAM, JOHN A.
APPLICANT: PINKO, CHRIS
APPLICANT: SHOWALTER, RICHARD
APPLICANT: PARAST, CAMRAN V.
APPLICANT: TEMPICZYK-RUSSEL, ANNA
APPLICANT: GEHRING, MICHAEL R.
APPLICANT: MROCKZKOWSKI, BARBARA
APPLICANT: KAN, CHEN-CHEN
APPLICANT: VILLAFRANCA, J. ERNEST
APPLICANT: APPERT, KRISTOF
TITLE OF INVENTION: MODIFICATIONS OF THE VEGF RECEPTOR-2 PROTEIN AND
METHODS OF USE
FILE REFERENCE: 0125-001605
CURRENT APPLICATION NUMBER: US/09/939,833
CURRENT FILING DATE: 2001-08-28
PRIOR APPLICATION NUMBER: US/09/390,326
PRIOR FILING DATE: 1999-09-07
NUMBER OF SEQ ID NOS: 12
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 6
LENGTH: 386
TYPE: PRT
ORGANISM: E. coli
US-09-939-833-6

Query Match
Best Local Similarity 33.3%; Score 5; DB 9; Length 386;
100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HLGP 6
|||||
DB 75 HLGP 79

```
RESULT 93
US-09-939-754-6
; Sequence 6, Application US/09939754
; Patent No. US20020051965A1
; GENERAL INFORMATION:
; APPLICANT: MCTIGUE, MICHELE A.
; APPLICANT: WICKERSHAM, JOHN A.
; APPLICANT: PINKO, CHRIS
; APPLICANT: SHOWALTER, RICHARD
; APPLICANT: PARAST, CAMRAN V.
; APPLICANT: TEMPCZYK-RUSSEL, ANNA
; APPLICANT: GEHRING, MICHAEL R.
; APPLICANT: MROCKOWSKI, BARBARA
; APPLICANT: KAN, CHEN-CHEN
; APPLICANT: VILLAFRANCA, J. ERNEST
; APPLICANT: APPELT, KRYSZTOF
; TITLE OF INVENTION: MODIFICATIONS OF THE VEGF RECEPTOR-2 PROTEIN AND
; FILE REFERENCE: 0125-0016US
; CURRENT APPLICATION NUMBER: US/09/939,754
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 09/390,326
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 386
; TYPE: PRT
; ORGANISM: E. coli
US-09-939-754-6

Query Match
Best Local Similarity 100.0%; Score 5; DB 10; Length 386;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HLGPB 6
DB 75 HLGPB 79

RESULT 94
US-09-939-832-6
; Sequence 6, Application US/09939832
; Patent No. US20020127538A1
; GENERAL INFORMATION:
; APPLICANT: MCTIGUE, MICHELE A.
; APPLICANT: WICKERSHAM, JOHN A.
; APPLICANT: PINKO, CHRIS
; APPLICANT: SHOWALTER, RICHARD
; APPLICANT: PARAST, CAMRAN V.
; APPLICANT: TEMPCZYK-RUSSEL, ANNA
; APPLICANT: GEHRING, MICHAEL R.
; APPLICANT: MROCKOWSKI, BARBARA
; APPLICANT: KAN, CHEN-CHEN
; APPLICANT: VILLAFRANCA, J. ERNEST
; APPLICANT: APPELT, KRYSZTOF
; TITLE OF INVENTION: MODIFICATIONS OF THE VEGF RECEPTOR-2 PROTEIN AND
; FILE REFERENCE: 0125-0016US
; CURRENT APPLICATION NUMBER: US/09/939,832
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 09/390,326
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 386
; TYPE: PRT
; ORGANISM: E. coli
US-09-939-832-6

Query Match
Best Local Similarity 100.0%; Score 5; DB 10; Length 386;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HLGPB 6
DB 75 HLGPB 79

Query Match
Best Local Similarity 100.0%; Score 5; DB 10; Length 386;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HLGPB 6
DB 75 HLGPB 79
```

```
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HLGPB 6
DB 75 HLGPB 79

RESULT 95
US-09-729-674-172
; Sequence 172, Application US/09729674
; Patent No. US20010039335A1
; GENERAL INFORMATION:
; APPLICANT: JACOBS, Kenneth
; APPLICANT: MCCOY, John M.
; APPLICANT: LAVAILLE, Edward R.
; APPLICANT: COLLINS-RACIE, Lisa A.
; APPLICANT: EVANS, Cheryl
; APPLICANT: MERBERG, David
; APPLICANT: TREACY, Maurice
; APPLICANT: AGOSTINO, Michael J.
; APPLICANT: STEININGER II, Robert J.
; APPLICANT: SPAUDLING, Vikki
; APPLICANT: MONY, Gordon G.
; APPLICANT: CLARK, Hilary
; APPLICANT: RECHTEL, Kim
; APPLICANT: GENETICS INSTITUTE, INC.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 6055-64X
; CURRENT APPLICATION NUMBER: US/09/729,674
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 09/539,330
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 283
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 172
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-729-674-172

Query Match
Best Local Similarity 100.0%; Score 5; DB 10; Length 420;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPEB 12
DB 320 STPEB 324

RESULT 96
US-09-742-954-14
; Sequence 14, Application US/09742954
; Patent No. US20010005749A1
; GENERAL INFORMATION:
; APPLICANT: CAHOON, Rebecca B.
; APPLICANT: FALCO, S. Carl
; APPLICANT: MORGANTE, Michele
; APPLICANT: RAFALSKI, J. Antoni
; APPLICANT: HITZ, William D.
; APPLICANT: KINNEY, Anthony J.
; TITLE OF INVENTION: Aromatic Amino Acid Catabolism Enzymes
; FILE REFERENCE: BB-1197
; CURRENT APPLICATION NUMBER: US/09/742,954
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/094,783
; PRIOR FILING DATE: JULY 31, 1998
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Glycine max
US-09-742-954-14
```

US-09-742-954-14

Query Match 33.3%; Score 5; DB 10; Length 421;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0;

QY 10 PESRA 14

DB 361 PESRA 365

RESULT 97

US-09-286-488-34
; Sequence 34, Application US/09286488
; Patent No. US20020169136A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS
; TITLE OF INVENTION: AND TREATMENT OF B. MICROTI INFECTION
; FILE REFERENCE: 210121.426C3
; CURRENT APPLICATION NUMBER: US/09/286,488
; CURRENT FILING DATE: 1999-04-05
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Babesia microti
US-09-286-488-34

Query Match 33.3%; Score 5; DB 9; Length 431;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0;

QY 8 STPES 12

DB 70 STPES 74

RESULT 98

US-09-737-178-34
; Sequence 34, Application US/09737178
; Patent No. US20010029295A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Homer, Mary
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS
; TITLE OF INVENTION: AND TREATMENT OF B. MICROTI INFECTION
; FILE REFERENCE: 210121.426C9
; CURRENT APPLICATION NUMBER: US/09/737,178
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Babesia microti
US-09-737-178-34

Query Match 33.3%; Score 5; DB 10; Length 431;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0;

QY 8 STPES 12

DB 70 STPES 74

RESULT 99

US-10-108-605-37
; Sequence 37, Application US/10108605
; Patent No. US20020160934A1
; GENERAL INFORMATION:
; APPLICANT: Broadus, Julie
; APPLICANT: Stam, Lynn
; APPLICANT: Bachmann, Jane
; APPLICANT: Kamdar, Kim
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCOD
; TITLE OF INVENTION: PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF
; FILE REFERENCE: 311338
; CURRENT APPLICATION NUMBER: US/10/108,605
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 09/761,142
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/176,418
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-108-605-37

Query Match 33.3%; Score 5; DB 9; Length 435;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0;

QY 11 ESRAA 15

DB 328 ESRAA 332

RESULT 100

US-09-815-242-5207
; Sequence 5207, Application US/09815242
; Patent No. US2002061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0


```
; SEQ ID NO 5207
; LENGTH: 439
; TYPE: PRF
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-5207
```

```
Query Match          33.3%; Score 5; DB 10; Length 439;
Best Local Similarity 100.0%; Pred. No. 4, 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      11 ESRAA 15
        |||||
Db       63 ESRAA 67
```

```
RESULT 101
US-09-738-626-4037
; Sequence 4037, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
```

```
; APPLICANT: MAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: OKADA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4037
; LENGTH: 453
; TYPE: PRF
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4037
```

```
Query Match          33.3%; Score 5; DB 9; Length 453;
Best Local Similarity 100.0%; Pred. No. 4, 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      11 ESRAA 15
        |||||
Db       281 ESRAA 285
```

```
RESULT 102
US-09-815-242-10237
; Sequence 10237, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; PROKARYOTES
; FILE REFERENCE: ELITRA.011A
```

```
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10237
; LENGTH: 471
; TYPE: PRF
; ORGANISM: Escherichia coli
US-09-815-242-10237
```

```
Query Match          33.3%; Score 5; DB 10; Length 471;
Best Local Similarity 100.0%; Pred. No. 4, 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      7 RSTPE 11
        |||||
Db       48 RSTPE 52
```

```
RESULT 103
US-09-815-242-13956
; Sequence 13956, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; PROKARYOTES
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13956
; LENGTH: 471
; TYPE: PRF
; ORGANISM: Salmonella typhi
; FEATURE:
; NAME/KEY: VARIANT
```

LOCATION: (1)...(471)
OTHER INFORMATION: Xaa = Any Amino Acid
US-09-815-242-13956

Query Match 33.3%; Score 5; DB 10; Length 471;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPE 11
|||||

DB 48 RSTPE 52

RESULT 104
US-09-796-753-158
Sequence 158, Application US/09796753
Publication No. US20030027998A1
GENERAL INFORMATION:
APPLICANT: McCarthy, Sean A.
TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
FILE REFERENCE: 7853-227-999
CURRENT APPLICATION NUMBER: US/09/796,753
CURRENT FILING DATE: 2001-03-01
PRIOR APPLICATION NUMBER: 09/183,175
PRIOR FILING DATE: 1998-10-30
PRIOR APPLICATION NUMBER: 09/223,094
PRIOR FILING DATE: 1998-12-30
PRIOR APPLICATION NUMBER: 09/223,546
PRIOR FILING DATE: 1998-12-30
PRIOR APPLICATION NUMBER: 09/224,246
PRIOR FILING DATE: 1998-12-30
PRIOR APPLICATION NUMBER: 09/259,388
PRIOR FILING DATE: 1999-02-26
PRIOR APPLICATION NUMBER: 60/122,458
PRIOR FILING DATE: 1999-03-01
PRIOR APPLICATION NUMBER: 09/312,359
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 09/336,536
PRIOR FILING DATE: 1999-06-18
PRIOR APPLICATION NUMBER: 09/342,687
PRIOR FILING DATE: 1999-06-29
PRIOR APPLICATION NUMBER: 09/345,464
PRIOR FILING DATE: 1999-06-30
PRIOR APPLICATION NUMBER: 09/365,164
PRIOR FILING DATE: 1999-07-30
PRIOR APPLICATION NUMBER: 09/399,723
PRIOR FILING DATE: 1999-09-20
PRIOR APPLICATION NUMBER: 09/409,634
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: 09/471,179
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: 09/474,071
PRIOR FILING DATE: 1999-12-29
PRIOR APPLICATION NUMBER: 09/474,072
PRIOR FILING DATE: 1999-12-29
PRIOR APPLICATION NUMBER: 09/514,010
PRIOR FILING DATE: 2000-02-25
PRIOR APPLICATION NUMBER: 09/516,745
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: 09/572,002
PRIOR FILING DATE: 2000-05-14
PRIOR APPLICATION NUMBER: 09/597,993
PRIOR FILING DATE: 2000-06-19
PRIOR APPLICATION NUMBER: 09/599,596
PRIOR FILING DATE: 2000-06-22
PRIOR APPLICATION NUMBER: 09/630,334
PRIOR FILING DATE: 2000-07-31
PRIOR APPLICATION NUMBER: 09/606,565
PRIOR FILING DATE: 2000-06-29
PRIOR APPLICATION NUMBER: 09/606,317
PRIOR FILING DATE: 2000-06-29
PRIOR APPLICATION NUMBER: 09/665,666
PRIOR FILING DATE: 2000-09-20

PRIOR APPLICATION NUMBER: 09/677,751
PRIOR FILING DATE: 2000-09-30
NUMBER OF SEQ ID NOS: 162
SEQ ID NO 158
LENGTH: 480
TYPE: PRT
ORGANISM: Homo sapiens
US-09-796-753-158

Query Match 33.3%; Score 5; DB 9; Length 480;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRRA 15
|||||

DB 248 ESRRA 252

RESULT 105
US-09-815-242-10999
Sequence 10999, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes In
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10999
LENGTH: 480
TYPE: PRT
ORGANISM: Haemophilus influenzae
US-09-815-242-10999

Query Match 33.3%; Score 5; DB 10; Length 480;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPE 11
|||||

DB 60 RSTPE 64

RESULT 106
US-09-739-254-154
Sequence 154, Application US/09739254
Patent No. US20010021700A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.

;; TITLE OF INVENTION: 49 Human Secreted Proteins
;; FILE REFERENCE: P2032P1
;; CURRENT APPLICATION NUMBER: US/09/739,254
;; CURRENT FILING DATE: 2000-12-19
;; EARLIER APPLICATION NUMBER: 09/511,554
;; EARLIER FILING DATE: 2000-02-23
;; EARLIER APPLICATION NUMBER: PCT/US99/19330
;; EARLIER FILING DATE: 1999-08-24
;; EARLIER APPLICATION NUMBER: 60/097,917
;; EARLIER FILING DATE: 1998-08-25
;; EARLIER APPLICATION NUMBER: 60/098,634
;; EARLIER FILING DATE: 1998-08-31
;; NUMBER OF SEQ ID NOS: 170
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 154
;; LENGTH: 483
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: SITE
;; LOCATION: (194)
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;; FEATURE:
;; NAME/KEY: SITE
;; LOCATION: (205)
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-739-254-154

Query Match 33.3%; Score 5; DB 10; Length 483;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGP 5
|||||
DB 437 SHLGP 441

RESULT 107
US-09-904-615-154
;; Sequence 154, Application US/09904615
;; Patent No. US20020026040A1
;; GENERAL INFORMATION:
;; APPLICANT: Rosen et al.
;; TITLE OF INVENTION: 49 Human Secreted Proteins
;; FILE REFERENCE: P2032P1
;; CURRENT APPLICATION NUMBER: US/09/904,615
;; CURRENT FILING DATE: 2001-07-16
;; PRIOR APPLICATION NUMBER: 09/511,554
;; PRIOR FILING DATE: 2000-02-23
;; PRIOR APPLICATION NUMBER: 60/097,917
;; PRIOR FILING DATE: 1998-08-25
;; PRIOR APPLICATION NUMBER: 60/098,634
;; PRIOR FILING DATE: 1998-08-31
;; NUMBER OF SEQ ID NOS: 170
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 154
;; LENGTH: 483
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: SITE
;; LOCATION: (194)
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;; NAME/KEY: SITE
;; LOCATION: (205)
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-904-615-154

Query Match 33.3%; Score 5; DB 10; Length 483;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGP 5

DB 437 SHLGP 441
|||||

RESULT 108
US-09-910-430-18
;; Sequence 18, Application US/09910430
;; Patent No. US20020127235A1
;; GENERAL INFORMATION:
;; APPLICANT: Godfroi, Edmond
;; APPLICANT: Bolien, Alex
;; APPLICANT: Lebouille, Gerard
;; TITLE OF INVENTION: IDENTIFICATION AND MOLECULAR CHARACTERIZATION OF
;; TITLE OF INVENTION: PROTEINS, EXPRESSED IN THE IXODES RICINUS SALIVARY
;; FILE REFERENCE: VAMM29_001CPI
;; CURRENT APPLICATION NUMBER: US/09/910,430
;; CURRENT FILING DATE: 2001-07-19
;; PRIOR APPLICATION NUMBER: PCT/BE00/00061
;; PRIOR FILING DATE: 2000-06-06
;; PRIOR APPLICATION NUMBER: GB 9913425.6
;; PRIOR FILING DATE: 1999-06-09
;; NUMBER OF SEQ ID NOS: 34
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 18
;; LENGTH: 488
;; TYPE: PRT
;; ORGANISM: Ixodes ricinus
US-09-910-430-18

Query Match 33.3%; Score 5; DB 10; Length 488;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRRA 15
|||||
DB 32 ESRRA 36

RESULT 109
US-10-029-180-106
;; Sequence 106, Application US/10029180
;; Publication No. US20020182708A1
;; GENERAL INFORMATION:
;; APPLICANT: Call, Brian M.
;; APPLICANT: Holtzman, Doug
;; APPLICANT: Madden, Kevin T.
;; APPLICANT: Milina, G. Todd
;; APPLICANT: Sherman, Amir
;; APPLICANT: Silva, Jeffrey C.
;; APPLICANT: Trueheart, Josh
;; APPLICANT: Zhang, Lixin
;; TITLE OF INVENTION: No. US20020182708A1e1 Regulators of Fungal Gene Expression
;; FILE REFERENCE: MTC-004
;; CURRENT APPLICATION NUMBER: US/10/029,180
;; CURRENT FILING DATE: 2001-12-22
;; PRIOR APPLICATION NUMBER: US 60/257,431
;; PRIOR FILING DATE: 2000-12-22
;; NUMBER OF SEQ ID NOS: 138
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 106
;; LENGTH: 491
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: fungal gene
US-10-029-180-106

Query Match 33.3%; Score 5; DB 9; Length 491;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESR 13

Db 100 TPESR 104

RESULT 110

US-09-997-900-2
; Sequence 2, Application US/09997900
; Patent No. US20020053098A1
; GENERAL INFORMATION:
; APPLICANT: Kakefuda, Genichi
; APPLICANT: Costello, Colleen
; APPLICANT: Sun, Ming
; APPLICANT: Hu, Weiming
; TITLE OF INVENTION: Genes and Vectors for Conferring Herbicide Resistance
; FILE REFERENCE: In Plants
; FILE REFERENCE: 043753/241148 (5849-20A)
; CURRENT APPLICATION NUMBER: US/09/997,900
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/106,239
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: 09/426,568
; PRIOR FILING DATE: 1999-10-22
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 491
; TYPE: PRT
; ORGANISM: Arabidopsis sp.
US-09-997-900-2

Query Match

Best Local Similarity 33.3%; Score 5; DB 10; Length 491;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRA 14
Db 177 PESRA 181

RESULT 111

US-09-815-242-13042
; Sequence 13042, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 13042
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-13042

Query Match 33.3%; Score 5; DB 10; Length 496;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ESRAA 15
Db 268 ESRAA 272

RESULT 112

US-09-900-237-20
; Sequence 20, Application US/09900237
; Patent No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cellulose Synthases
; FILE REFERENCE: BR1170 US CIP
; CURRENT APPLICATION NUMBER: US/09/900,237
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092,844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/15871
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/720383
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 20
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (88)
US-09-900-237-20

Query Match 33.3%; Score 5; DB 10; Length 506;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STRES 12
Db 169 STRES 173

RESULT 113

US-09-996-015-45
; Sequence 45, Application US/09996015
; Publication No. US20030032166A1
; GENERAL INFORMATION:
; APPLICANT: Quinn, Kerry E.
; APPLICANT: Pena, Carol A. E.
; APPLICANT: Li, Li
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Lette, Mario W.
; TITLE OF INVENTION: Aortic Carboxypeptidase-Like Proteins and Nucleic Acids
; TITLE OF INVENTION: Encoding Same
; FILE REFERENCE: 15966-581 CIP
; CURRENT APPLICATION NUMBER: US/09/996,015
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 09/641,741
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 60/159,613
; PRIOR FILING DATE: 1999-10-14
; PRIOR APPLICATION NUMBER: 60/175,534
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 60/224,086

;; PRIOR FILING DATE: 2000-08-09
;; NUMBER OF SEQ ID NOS: 47
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 45
;; LENGTH: 510
;; TYPE: PRF
;; ORGANISM: Homo sapiens
US-09-996-015-45

Query Match 33.3%; Score 5; DB 9; Length 510;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LCPHR 7
DB 140 LCPHR 144

RESULT 114
US-10-028-072-210
; Sequence 210, Application US/10028072
; Publication No. US20030004311A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang
; TITLE OF INVENTION:
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/028,072
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059836
; PRIOR FILING DATE: 1997-09-24
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062285
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062814

;; PRIOR FILING DATE: 1997-10-24
;; PRIOR APPLICATION NUMBER: 60/062816
;; PRIOR FILING DATE: 1997-10-24
;; PRIOR APPLICATION NUMBER: 60/063045
;; PRIOR FILING DATE: 1997-10-24
;; PRIOR APPLICATION NUMBER: 60/063082
;; PRIOR FILING DATE: 1997-10-31
;; PRIOR APPLICATION NUMBER: 60/063127
;; PRIOR FILING DATE: 1997-10-24
;; PRIOR APPLICATION NUMBER: 60/063327
;; PRIOR FILING DATE: 1997-10-27
;; PRIOR APPLICATION NUMBER: 60/063329
;; PRIOR FILING DATE: 1997-10-27
;; PRIOR APPLICATION NUMBER: 60/063550
;; PRIOR FILING DATE: 1997-10-28
;; PRIOR APPLICATION NUMBER: 60/063561
;; PRIOR FILING DATE: 1997-10-28
;; PRIOR APPLICATION NUMBER: 60/063704
;; PRIOR FILING DATE: 1997-10-29
;; PRIOR APPLICATION NUMBER: 60/063733
;; PRIOR FILING DATE: 1997-10-29
;; PRIOR APPLICATION NUMBER: 60/063735
;; PRIOR FILING DATE: 1997-10-29
;; PRIOR APPLICATION NUMBER: 60/063738
;; PRIOR FILING DATE: 1997-10-29
;; PRIOR APPLICATION NUMBER: 60/063755
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/064248
;; PRIOR FILING DATE: 1997-11-03
;; PRIOR APPLICATION NUMBER: 60/064809
;; PRIOR FILING DATE: 1997-11-07
;; PRIOR APPLICATION NUMBER: 60/065186
;; PRIOR FILING DATE: 1997-11-12
;; PRIOR APPLICATION NUMBER: 60/065846
;; PRIOR FILING DATE: 1997-11-17
;; PRIOR APPLICATION NUMBER: 60/066364
;; PRIOR FILING DATE: 1997-11-21
;; PRIOR APPLICATION NUMBER: 60/066453
;; PRIOR FILING DATE: 1997-11-24
;; PRIOR APPLICATION NUMBER: 60/066511
;; PRIOR FILING DATE: 1997-11-24
;; PRIOR APPLICATION NUMBER: 60/066770
;; PRIOR FILING DATE: 1997-11-24
;; PRIOR APPLICATION NUMBER: 60/069212
;; PRIOR FILING DATE: 1997-12-11
;; PRIOR APPLICATION NUMBER: 60/069278
;; PRIOR FILING DATE: 1997-12-11
;; PRIOR APPLICATION NUMBER: 60/069334
;; PRIOR FILING DATE: 1997-12-11
;; PRIOR APPLICATION NUMBER: 60/069694
;; PRIOR FILING DATE: 1997-12-16
;; PRIOR APPLICATION NUMBER: 60/072320
;; PRIOR FILING DATE: 1998-01-23
;; PRIOR APPLICATION NUMBER: 60/073612
;; PRIOR FILING DATE: 1998-02-04
;; PRIOR APPLICATION NUMBER: 60/074086
;; PRIOR FILING DATE: 1998-02-09
;; PRIOR APPLICATION NUMBER: 60/074092
;; PRIOR FILING DATE: 1998-02-09
;; PRIOR APPLICATION NUMBER: 60/077791
;; PRIOR FILING DATE: 1998-03-12
;; PRIOR APPLICATION NUMBER: 60/078910
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/079234
;; PRIOR FILING DATE: 1998-03-25
;; PRIOR APPLICATION NUMBER: 60/079663
;; PRIOR FILING DATE: 1998-02-27
;; PRIOR APPLICATION NUMBER: 60/079728
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/080165
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/081203
;; PRIOR FILING DATE: 1998-04-09

;; PRIOR APPLICATION NUMBER: 60/081229
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081695
;; PRIOR FILING DATE: 1998-04-14
;; PRIOR APPLICATION NUMBER: 60/081817
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081818
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/082999
;; PRIOR FILING DATE: 1998-04-24
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/083545
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084627
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084637
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085149
;; PRIOR FILING DATE: 1998-05-12
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085338
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085339
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085697
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085704
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/086414
;; PRIOR FILING DATE: 1998-05-22
;; PRIOR APPLICATION NUMBER: 60/086430
;; PRIOR FILING DATE: 1998-05-22
;; PRIOR APPLICATION NUMBER: 60/087106
;; PRIOR FILING DATE: 1998-05-28
;; PRIOR APPLICATION NUMBER: 60/088026
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088730
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088741
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088810
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088858
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089599
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089947
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090538
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091982

;; PRIOR FILING DATE: 1998-07-07
Query Match 33.3%; Score 5; DB 9; Length 519;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 GPHRS 8
DB 69 GPHRS 73
RESULT 115
US-10-121-049-210
; Sequence 210, Application US/10121049
; Publication No. US2003002239A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C17
; CURRENT FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: US/10/121,049
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 210
; LENGTH: 519
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-121-049-210
Query Match 33.3%; Score 5; DB 9; Length 519;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 GPHRS 8
DB 69 GPHRS 73
RESULT 116
US-10-123-904-210
; Sequence 210, Application US/10123904
; Publication No. US20030022328A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.

```

; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C54
; CURRENT APPLICATION NUMBER: US/10/123,904
; CURRENT FILING DATE: 2002-04-16
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 210
; LENGTH: 519
; TYPE: PRT
; ORGANISM: Homo Saplen
US-10-123-904-210

Query Match      33.3%; Score 5; DB 9; Length 519;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 GPHRS 8
Db      69 GPHRS 73

RESULT 117
US-10-140-470-210
; Sequence 210, Application US/10140470
; Publication No. US20030022331A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geriltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C160
; CURRENT APPLICATION NUMBER: US/10/140,470
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 210
; LENGTH: 519
; TYPE: PRT
; ORGANISM: Homo Saplen
US-10-140-470-210

Query Match      33.3%; Score 5; DB 9; Length 519;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 GPHRS 8
Db      69 GPHRS 73

RESULT 118
US-10-175-746-210
; Sequence 210, Application US/10175746
```

```

; Publication No. US200300227270A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geriltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C382
; CURRENT APPLICATION NUMBER: US/10/176,918
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 210
; LENGTH: 519
; TYPE: PRT
US-10-175-746-210

Query Match      33.3%; Score 5; DB 9; Length 519;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 GPHRS 8
Db      69 GPHRS 73

RESULT 119
US-10-176-918-210
; Sequence 210, Application US/10176918
; Publication No. US20030022725A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geriltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C382
; CURRENT APPLICATION NUMBER: US/10/176,918
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 210
; LENGTH: 519
; TYPE: PRT
US-10-176-918-210
```

```
; ORGANISM: Homo Sapien
US-10-176-918-210

Query Match
Best Local Similarity 33.3%; Score 5; DB 9; Length 519;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GPHRS 8
    |||||
Db 69 GPHRS 73

RESULT 120
US-10-176-921-210
; Sequence 210, Application US/10176921
; Publication No. US20030027276A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C288
; CURRENT APPLICATION NUMBER: US/10/176,921
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 210
; LENGTH: 519
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-921-210

Query Match
Best Local Similarity 33.3%; Score 5; DB 9; Length 519;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GPHRS 8
    |||||
Db 69 GPHRS 73

RESULT 121
US-10-137-865-210
; Sequence 210, Application US/10137865
; Publication No. US20030032155A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
```

```
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C154
; CURRENT APPLICATION NUMBER: US/10/137,865
; CURRENT FILING DATE: 2002-05-03
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 210
; LENGTH: 519
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-137-865-210

Query Match
Best Local Similarity 33.3%; Score 5; DB 9; Length 519;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GPHRS 8
    |||||
Db 69 GPHRS 73

RESULT 122
US-10-140-474-210
; Sequence 210, Application US/10140474
; Publication No. US20030032156A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C162
; CURRENT APPLICATION NUMBER: US/10/140,474
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 210
; LENGTH: 519
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-474-210

Query Match
Best Local Similarity 33.3%; Score 5; DB 9; Length 519;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GPHRS 8
    |||||
Db 69 GPHRS 73

RESULT 123
US-10-142-431-210
```



```
; Sequence 210, Application US/10142431
; Publication No. US20030036179A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Collin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330RIC251
; CURRENT APPLICATION NUMBER: US/10/142,431
; CURRENT FILING DATE: 2002-05-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 210
; LENGTH: 519
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-142-431-210
```

```
Query Match          33.3%; Score 5; DB 9; Length 519;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 4 GPHRS 8
Db 69 GPHRS 73
```

```
RESULT 124
US-10-143-114-210
; Sequence 210, Application US/10143114
; Publication No. US20030036180A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Collin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330RIC211
; CURRENT APPLICATION NUMBER: US/10/143,114
; CURRENT FILING DATE: 2002-05-09
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 210
; LENGTH: 519
```

```
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-143-114-210
```

```
Query Match          33.3%; Score 5; DB 9; Length 519;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 4 GPHRS 8
Db 69 GPHRS 73
```

```
RESULT 125
US-10-140-002-210
; Sequence 210, Application US/10140002
; Publication No. US20030037623A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Collin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330RIC59
; CURRENT APPLICATION NUMBER: US/10/140,002
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 210
; LENGTH: 519
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-002-210
```

```
Query Match          33.3%; Score 5; DB 9; Length 519;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 4 GPHRS 8
Db 69 GPHRS 73
```

```
RESULT 126
US-10-001-843-133
; Sequence 133, Application US/10001843
; Patent No. US20020132255A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Hervey
; APPLICANT: Cafferty, Robert
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; APPLICANT: Turner, Leah
; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific Genes an
; FILE REFERENCE: DEX-0267
; CURRENT APPLICATION NUMBER: US/10/001,843
; CURRENT FILING DATE: 2001-11-20
```

```

; PRIOR APPLICATION NUMBER: 60/249,992
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 218
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 133
; LENGTH: 519
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-001-843-133

Query Match          33.3%; Score 5; DB 12; Length 519;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GPHRS 8
   |||||
Db 69 GPHRS 73

RESULT 127
US-09-996-015-6
; Sequence 6, Application US/09996015
; Publication No. US20030032166A1
; GENERAL INFORMATION:
; APPLICANT: Oulna, Kerry E.
; APPLICANT: Pena, Carol A. E.
; APPLICANT: Li, Li
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Leite, Mario W.
; TITLE OF INVENTION: Aortic Carboxypeptidase-Like Proteins and Nucleic Acids
; FILE REFERENCE: 15966-581 CIP
; CURRENT APPLICATION NUMBER: US/09/996,015
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 09/641,741
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 60/159,613
; PRIOR FILING DATE: 1999-10-14
; PRIOR APPLICATION NUMBER: 60/175,534
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 60/224,086
; PRIOR FILING DATE: 2000-08-09
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 574
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-996-015-6

Query Match          33.3%; Score 5; DB 9; Length 574;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPHR 7
   |||||
Db 140 LGPHR 144

RESULT 128
US-09-782-980-78
; Sequence 78, Application US/09782980
; Patent No. US20020072089A1
; GENERAL INFORMATION:
; APPLICANT: Rhododoust, Mehran M.
; APPLICANT: Macbeth, Kyle J.
; APPLICANT: Busfield, Samantha J.
; APPLICANT: McCarthy, Sean A.
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Gu, Wei
; APPLICANT: White, David
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL ITALY, LOR-2, STRIFE, TRASH, BDSF, LRSF, AND
```

```

; TITLE OF INVENTION: STMST PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: MNI-121CP
; CURRENT APPLICATION NUMBER: US/09/782,980
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: PCT/US00/02125
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: 09/448,076
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: 09/276,400
; PRIOR FILING DATE: 1999-03-25
; PRIOR APPLICATION NUMBER: 60/117,580
; PRIOR FILING DATE: 1999-01-27
; PRIOR APPLICATION NUMBER: 09/014,195
; PRIOR FILING DATE: 1998-01-27
; PRIOR APPLICATION NUMBER: 09/014,348
; PRIOR FILING DATE: 1998-01-27
; PRIOR APPLICATION NUMBER: 09/086,892
; PRIOR FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 09/296,208
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 09/063,950
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 09/561,381
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 09/561,810
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 09/087,121
; PRIOR FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 09/672,721
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 09/049,799
; PRIOR FILING DATE: 1998-03-27
; NUMBER OF SEQ ID NOS: 176
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 78
; LENGTH: 588
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-980-78

Query Match          33.3%; Score 5; DB 10; Length 588;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRA 14
   |||||
Db 543 PESRA 547

RESULT 129
US-09-738-626-5305
; Sequence 5305, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKTO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
```

PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 5305
LENGTH: 608
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-5305

Query Match 33.3%; Score 5; DB 9; Length 608;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
|||||
DB 171 RSTPE 175

RESULT 130
US-09-738-626-6059
Sequence 6059, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OKHAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OKAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 6059
LENGTH: 630
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-6059

Query Match 33.3%; Score 5; DB 9; Length 630;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPE 12
|||||
DB 163 STPE 167

RESULT 131
US-10-118-328-4
Sequence 4, Application US/10118328
Patent No. US20020169289A1
GENERAL INFORMATION:
APPLICANT: YAN, Chunhua et al.
TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
TITLE OF INVENTION: PROTEINS, AND USES THEREOF
FILE REFERENCE: C1001220
CURRENT APPLICATION NUMBER: US/10/118,328
CURRENT FILING DATE: 2002-04-09

PRIOR APPLICATION NUMBER: 60/282,460
PRIOR FILING DATE: 2001-04-10
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 709
TYPE: PRT
ORGANISM: Mus musculus
US-10-118-328-4

Query Match 33.3%; Score 5; DB 9; Length 709;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
|||||
DB 38 RSTPE 42

RESULT 132
US-09-923-444A-2
Sequence 2, Application US/09923444A
Patent No. US20020015980A1
GENERAL INFORMATION:
APPLICANT: LI, YI
TITLE OF INVENTION: NEUROTRANSMITTER TRANSPORTER
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: MD
COUNTRY: US
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/923,444A
FILING DATE: 08-Aug-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/062,815
FILING DATE: 199-12-20
ATTORNEY/AGENT INFORMATION:
NAME: Michele M. Wales
REGISTRATION NUMBER: 43,975
REFERENCE/DOCKET NUMBER: P1116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 727 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-923-444A-2

Query Match 33.3%; Score 5; DB 10; Length 727;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 STPE 12
|||||
DB 721 STPE 725

RESULT 133

```
US-09-862-027-42
; Sequence 42, Application US/09862027
; Patent No. US20020142428A1
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; TITLE OF INVENTION: No. US20020142428A1 Kinases and Uses Thereof
; FILE REFERENCE: 35800/234862
; CURRENT APPLICATION NUMBER: US/09/862,027
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 09/345,473
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 733
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-862-027-42

Query Match
; 33.3%; Score 5; DB 9; Length 733;
; Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRA 14
; 11111
Db 567 PESRA 571

RESULT 134
US-10-174-590-458
; Sequence 458, Application US/10174590
; Publication No. US20030008352A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C42
; CURRENT APPLICATION NUMBER: US/10/174,590
; CURRENT FILING DATE: 2002-06-18
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-590-458

Query Match
; 33.3%; Score 5; DB 9; Length 734;
; Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPHR 7
; 11111
Db 140 LGPHR 144

RESULT 135
US-10-176-758-458
; Sequence 458, Application US/10176758
; Publication No. US20030008353A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
```

```
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C104
; CURRENT APPLICATION NUMBER: US/10/176,758
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-758-458

Query Match
; 33.3%; Score 5; DB 9; Length 734;
; Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPHR 7
; 11111
Db 140 LGPHR 144

RESULT 136
US-10-175-737-458
; Sequence 458, Application US/10175737
; Publication No. US20030013153A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C50
; CURRENT APPLICATION NUMBER: US/10/175,737
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-737-458

Query Match
; 33.3%; Score 5; DB 9; Length 734;
; Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPHR 7
; 11111
Db 140 LGPHR 144

RESULT 137
US-10-173-706-458
; Sequence 458, Application US/10173706
; Publication No. US2003002293A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C7
; CURRENT APPLICATION NUMBER: US/10/173,706
; PRIOR FILING DATE: 2002-06-17
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Saplen
US-10-173-706-458

Query Match          33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      3 LGPFR 7
      |||||
Db      140 LGPFR 144

RESULT 138
US-10-175-738-458
; Sequence 458, Application US/10175738
; Publication No. US20030022294A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C45
; CURRENT APPLICATION NUMBER: US/10/175,738
; PRIOR FILING DATE: 2002-06-19
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Saplen
US-10-175-738-458

Query Match          33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      3 LGPFR 7
      |||||
Db      140 LGPFR 144

RESULT 139
```

```

US-10-175-752-458
; Sequence 458, Application US/10175752
; Publication No. US20030022295A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C60
; CURRENT APPLICATION NUMBER: US/10/175,752
; PRIOR FILING DATE: 2002-06-19
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Saplen
US-10-175-752-458

Query Match          33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      3 LGPFR 7
      |||||
Db      140 LGPFR 144

RESULT 140
US-10-176-482-458
; Sequence 458, Application US/10176482
; Publication No. US20030022296A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C70
; CURRENT APPLICATION NUMBER: US/10/176,482
; PRIOR FILING DATE: 2002-06-20
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Saplen
US-10-176-482-458

Query Match          33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      3 LGPFR 7
      |||||
Db      140 LGPFR 144
```

RESULT 141
US-10-176-757-458
; Sequence 458, Application US/10176757
; Publication No. US20030022297A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C86
; CURRENT APPLICATION NUMBER: US/10/176,757
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-757-458

Query Match 33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 3 LGPHR 7
Db 140 LGPHR 144

RESULT 142
US-10-176-913-458
; Sequence 458, Application US/10176913
; Publication No. US20030022298A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C66
; CURRENT APPLICATION NUMBER: US/10/176,913
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-913-458

Query Match 33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 LGPHR 7
Db 140 LGPHR 144

RESULT 143
US-10-180-552-458
; Sequence 458, Application US/10180552
; Publication No. US20030022300A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C153
; CURRENT APPLICATION NUMBER: US/10/180,552
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-180-552-458

Query Match 33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 3 LGPHR 7
Db 140 LGPHR 144

RESULT 144
US-10-180-557-458
; Sequence 458, Application US/10180557
; Publication No. US20030022301A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C147
; CURRENT APPLICATION NUMBER: US/10/180,557
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-180-557-458

Query Match 33.3%; Score 5; DB 9; Length 734;

Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPFR 7
|||||

Db 140 LGPFR 144

RESULT 145

US-10-173-700-458

; Sequence 458, Application US/10173700

; Publication No. US20030027262A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430R1C14

; CURRENT APPLICATION NUMBER: US/10/173,700

; FILE REFERENCE: P3430R1C14

; CURRENT FILING DATE: 2002-06-17

; Prior Application removed - See file wrapper or Palm

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 458

; LENGTH: 734

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-173-700-458

Query Match 33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPFR 7
|||||

Db 140 LGPFR 144

RESULT 146

US-10-174-572-458

; Sequence 458, Application US/10174572

; Publication No. US20030027263A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430R1C40

; CURRENT APPLICATION NUMBER: US/10/174,572

; FILE REFERENCE: P3430R1C40

; CURRENT FILING DATE: 2002-06-18

; Prior Application removed - See file wrapper or Palm

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 458

; LENGTH: 734

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-174-572-458

Query Match 33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPFR 7
|||||

Db 140 LGPFR 144

RESULT 147

US-10-174-579-458

; Sequence 458, Application US/10174579

; Publication No. US20030027264A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430R1C31

; CURRENT APPLICATION NUMBER: US/10/174,579

; FILE REFERENCE: P3430R1C31

; CURRENT FILING DATE: 2002-06-18

; Prior Application removed - See file wrapper or Palm

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 458

; LENGTH: 734

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-174-579-458

Query Match 33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPFR 7
|||||

Db 140 LGPFR 144

RESULT 148

US-10-174-582-458

; Sequence 458, Application US/10174582

; Publication No. US20030027265A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430R1C36

; CURRENT APPLICATION NUMBER: US/10/174,582

; FILE REFERENCE: P3430R1C36

; CURRENT FILING DATE: 2002-06-18

; Prior Application removed - See file wrapper or Palm

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 458

LENGTH: 734
TYPE: PRT
ORGANISM: Homo Sapien
US-10-174-582-458

Query Match
Best Local Similarity 100.0%; Score 5; DB 9; Length 734;
Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHR 7
|||||
DB 140 LGPHR 144

RESULT 149
US-10-174-588-458
Sequence 458, Application US/10174588
Publication No. US20030027266A1
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3430R1C28
CURRENT APPLICATION NUMBER: US/10/174, 588
CURRENT FILING DATE: 2002-06-18
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 458
LENGTH: 734
TYPE: PRT
ORGANISM: Homo Sapien
US-10-174-588-458

Query Match
Best Local Similarity 100.0%; Score 5; DB 9; Length 734;
Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHR 7
|||||
DB 140 LGPHR 144

RESULT 150
US-10-175-739-458
Sequence 458, Application US/10175739
Publication No. US20030027267A1
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3430R1C46
CURRENT APPLICATION NUMBER: US/10/175, 739
CURRENT FILING DATE: 2002-06-19

Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 458
LENGTH: 734
TYPE: PRT
ORGANISM: Homo Sapien
US-10-175-739-458

Query Match
Best Local Similarity 100.0%; Score 5; DB 9; Length 734;
Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHR 7
|||||
DB 140 LGPHR 144

RESULT 151
US-10-175-740-458
Sequence 458, Application US/10175740
Publication No. US20030027268A1
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3430R1C61
CURRENT APPLICATION NUMBER: US/10/175, 740
CURRENT FILING DATE: 2002-06-18
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 458
LENGTH: 734
TYPE: PRT
ORGANISM: Homo Sapien
US-10-175-740-458

Query Match
Best Local Similarity 100.0%; Score 5; DB 9; Length 734;
Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHR 7
|||||
DB 140 LGPHR 144

RESULT 152
US-10-175-743-458
Sequence 458, Application US/10175743
Publication No. US20030027269A1
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME

FILE REFERENCE: P3430RIC52
CURRENT APPLICATION NUMBER: US/10/175,743
CURRENT FILING DATE: 2002-06-16
PRIOR APPLICATION NUMBER: 10/052586
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063540
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063564
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063734
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063870
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066120
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066466
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066772
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069335
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069425
PRIOR FILING DATE: 1997-12-12
PRIOR APPLICATION NUMBER: 60/069870
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/068017
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086023
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/086392
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/086486
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087098
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087208
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217

```
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088326
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088655
;; PRIOR FILING DATE: 1998-06-09
;; PRIOR APPLICATION NUMBER: 60/088722
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088738
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088740
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088811
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088824
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088825
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088826
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088861
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088863
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088876
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089090
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089105
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089512
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
```

```
Query Match          33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 3 LGPFR 7
      |||||
Db 140 LGPFR 144
```

```
RESULT 153
US-10-176-488-458
; Sequence 458, Application US/10176488
; Publication No. US20030027271A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C119
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
```

```
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-10-176-488-458
```

```
Query Match          33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 3 LGPFR 7
      |||||
Db 140 LGPFR 144
```

```
RESULT 154
US-10-176-492-458
; Sequence 458, Application US/10176492
; Publication No. US20030027272A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C107
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-492-458
```

```
Query Match          33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 3 LGPFR 7
      |||||
Db 140 LGPFR 144
```

```
RESULT 155
US-10-176-747-458
; Sequence 458, Application US/10176747
; Publication No. US20030027273A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C92
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
```

; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Saplen
US-10-176-747-458

Query Match 33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPFR 7
|||||
DB 140 LGPFR 144

RESULT 156
US-10-176-750-458
; Sequence 458, Application US/10176750
; Publication No. US20030027274A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C103
; CURRENT APPLICATION NUMBER: US/10/176,750
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Saplen
US-10-176-750-458

Query Match 33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPFR 7
|||||
DB 140 LGPFR 144

RESULT 157
US-10-176-985-458
; Sequence 458, Application US/10176985
; Publication No. US20030027277A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C99

; CURRENT APPLICATION NUMBER: US/10/176,985
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Saplen
US-10-176-985-458

Query Match 33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPFR 7
|||||
DB 140 LGPFR 144

RESULT 158
US-10-176-987-458
; Sequence 458, Application US/10176987
; Publication No. US20030027278A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C93
; CURRENT APPLICATION NUMBER: US/10/176,987
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Saplen
US-10-176-987-458

Query Match 33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPFR 7
|||||
DB 140 LGPFR 144

RESULT 159
US-10-176-991-458
; Sequence 458, Application US/10176991
; Publication No. US20030027324A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin

```
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C122
; CURRENT APPLICATION NUMBER: US/10/176,991
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-991-458

Query Match      33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3 LGPFR 7
        |||||
Db      140 LGPFR 144

RESULT 160
US-10-176-992-458
; Sequence 458, Application US/10176992
; Publication No. US20030027279A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C100
; CURRENT APPLICATION NUMBER: US/10/176,992
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-992-458
```

```
Query Match      33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3 LGPFR 7
        |||||
Db      140 LGPFR 144

RESULT 161
US-10-176-993-458
; Sequence 458, Application US/10176993
; Publication No. US20030027280A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
```

```
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C89
; CURRENT APPLICATION NUMBER: US/10/176,993
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-993-458

Query Match      33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3 LGPFR 7
        |||||
Db      140 LGPFR 144

RESULT 162
US-10-184-658-458
; Sequence 458, Application US/10184658
; Publication No. US20030027281A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C228
; CURRENT APPLICATION NUMBER: US/10/184,658
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-184-658-458
```

```
Query Match      33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3 LGPFR 7
        |||||
Db      140 LGPFR 144

RESULT 163
US-09-996-015-2
; Sequence 2, Application US/09996015
; Publication No. US20030032166A1
; GENERAL INFORMATION:
; APPLICANT: Quinn, Kerry E.
; APPLICANT: Pena, Carol A. E.
; APPLICANT: Li, Li
; APPLICANT: Spaderina, Steven K.
; APPLICANT: Leite, Mario W.
```

```
; TITLE OF INVENTION: Aortic Carboxypeptidase-Like Proteins and Nucleic Acids
; FILE REFERENCE: 15966-581 CIP
; CURRENT APPLICATION NUMBER: US/09/996,015
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 09/641,741
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 60/159,613
; PRIOR FILING DATE: 1999-10-14
; PRIOR APPLICATION NUMBER: 60/175,534
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 60/224,086
; PRIOR FILING DATE: 2000-08-09
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-996-015-2
```

```
Query Match          33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      3 LGPFR 7
        |||||
Db      140 LGPFR 144
```

```
RESULT 164
US-10-173-695-458
; Sequence 458, Application US/10173695
; Publication No. US20030032101A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C3
; CURRENT APPLICATION NUMBER: US/10/173,695
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-173-695-458
```

```
Query Match          33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      3 LGPFR 7
        |||||
Db      140 LGPFR 144
```

```
RESULT 165
US-10-173-697-458
; Sequence 458, Application US/10173697
; Publication No. US20030032102A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C5
; CURRENT APPLICATION NUMBER: US/10/173,697
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-173-697-458
```

```
Query Match          33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      3 LGPFR 7
        |||||
Db      140 LGPFR 144
```

```
RESULT 166
US-10-173-705-458
; Sequence 458, Application US/10173705
; Publication No. US20030032103A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C18
; CURRENT APPLICATION NUMBER: US/10/173,705
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-173-705-458
```

```
Query Match          33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      3 LGPFR 7
        |||||
Db      140 LGPFR 144
```

```
RESULT 167
US-10-174-576-458
```

```
Sequence 458, Application US/10174576
; Publication No. US20030032104A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C23
; CURRENT APPLICATION NUMBER: US/10/174,576
; CURRENT FILING DATE: 2002-06-18
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-576-458
```

```
Query Match          33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      3 LGPHR 7
        |||||
Db      140 LGPHR 144
```

```
RESULT 168
US-10-174-585-458
; Sequence 458, Application US/10174585
; Publication No. US20030032105A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C37
; CURRENT APPLICATION NUMBER: US/10/174,585
; CURRENT FILING DATE: 2002-06-18
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-585-458
```

```
Query Match          33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      3 LGPHR 7
        |||||
Db      140 LGPHR 144
```

```
RESULT 169
US-10-174-586-458
; Sequence 458, Application US/10174586
; Publication No. US20030032106A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C24
; CURRENT APPLICATION NUMBER: US/10/174,586
; CURRENT FILING DATE: 2002-06-18
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-586-458
```

```
Query Match          33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      3 LGPHR 7
        |||||
Db      140 LGPHR 144
```

```
RESULT 170
US-10-175-747-458
; Sequence 458, Application US/10175747
; Publication No. US20030032107A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C44
; CURRENT APPLICATION NUMBER: US/10/175,747
; CURRENT FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
```



```
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088863
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089090
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
```

```
Query Match      33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      3 LGPR 7
        |||||
Db      140 LGPR 144
```

```
RESULT 171
US-10-176-481-458
; Sequence 458, Application US/10176481
; Publication No. US20030032108A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Uian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C98
; CURRENT APPLICATION NUMBER: US/10/176,481
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-481-458
```

```
Query Match      33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      3 LGPR 7
        |||||
Db      140 LGPR 144
```

```
RESULT 172
US-10-176-485-458
; Sequence 458, Application US/10176485
```

```
; Publication No. US20030032109A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Uian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C78
; CURRENT APPLICATION NUMBER: US/10/176,485
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-485-458
```

```
Query Match      33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      3 LGPR 7
        |||||
Db      140 LGPR 144
```

```
RESULT 173
US-10-176-487-458
; Sequence 458, Application US/10176487
; Publication No. US20030032110A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Uian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C74
; CURRENT APPLICATION NUMBER: US/10/176,487
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-487-458
```

```
Query Match      33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      3 LGPR 7
        |||||
Db      140 LGPR 144
```


RESULT 174
US-10-176-493-458
; Sequence 458, Application US/10176493
; Publication No. US2003003211A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C72
; CURRENT APPLICATION NUMBER: US/10/176,493
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-493-458

Query Match 33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPFR 7
|||||
DB 140 LGPFR 144

RESULT 175
US-10-176-756-458
; Sequence 458, Application US/10176756
; Publication No. US2003003211A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C109
; CURRENT APPLICATION NUMBER: US/10/176,756
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-756-458

Query Match 33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 3 LGPFR 7
|||||

DB 140 LGPFR 144

RESULT 176
US-10-176-911-458
; Sequence 458, Application US/10176911
; Publication No. US2003003211A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C75
; CURRENT APPLICATION NUMBER: US/10/176,911
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-911-458

Query Match 33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPFR 7
|||||
DB 140 LGPFR 144

RESULT 177
US-10-176-919-458
; Sequence 458, Application US/10176919
; Publication No. US2003003211A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C63
; CURRENT APPLICATION NUMBER: US/10/176,919
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-919-458

Query Match 33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPFR 7
|||||
Db 140 LGPFR 144

RESULT 178

US-10-176-925-458
; Sequence 458, Application US/10176925
; Publication No. US20030032115A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C94
; CURRENT APPLICATION NUMBER: US/10/176,925
; CURRENT FILING DATE: 2002-06-21
; Prior application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-925-458

Query Match 33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPFR 7
|||||
Db 140 LGPFR 144

RESULT 179

US-10-176-978-458
; Sequence 458, Application US/10176978
; Publication No. US20030032116A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C116
; CURRENT APPLICATION NUMBER: US/10/176,978
; CURRENT FILING DATE: 2002-06-21
; Prior application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-978-458

Query Match 33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPFR 7
|||||
Db 140 LGPFR 144

RESULT 180

US-10-179-510-458
; Sequence 458, Application US/10179510
; Publication No. US20030032117A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C138
; CURRENT APPLICATION NUMBER: US/10/179,510
; CURRENT FILING DATE: 2002-06-24
; Prior application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-179-510-458

Query Match 33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPFR 7
|||||
Db 140 LGPFR 144

RESULT 181

US-10-180-543-458
; Sequence 458, Application US/10180543
; Publication No. US20030032118A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C161
; CURRENT APPLICATION NUMBER: US/10/180,543
; CURRENT FILING DATE: 2002-06-25
; Prior application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT

ORGANISM: Homo Sapien
US-10-180-543-458

Query Match 33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPFR 7
|||||
Db 140 LGPFR 144

RESULT 182
US-10-180-544-458

Sequence 458, Application US/10180544
Publication No. US20030032119A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C150
CURRENT APPLICATION NUMBER: US/10/180,544
CURRENT FILING DATE: 2002-06-25
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 458
LENGTH: 734
TYPE: PRT
ORGANISM: Homo Sapien
US-10-180-544-458

Query Match 33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPFR 7
|||||
Db 140 LGPFR 144

RESULT 183
US-10-180-546-458

Sequence 458, Application US/10180546
Publication No. US20030032120A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C156
CURRENT APPLICATION NUMBER: US/10/180,546
CURRENT FILING DATE: 2002-06-25
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612

SEQ ID NO 458
LENGTH: 734
TYPE: PRT

ORGANISM: Homo Sapien
US-10-180-546-458

Query Match 33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPFR 7
|||||
Db 140 LGPFR 144

RESULT 184
US-10-180-547-458

Sequence 458, Application US/10180547
Publication No. US20030032121A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C157
CURRENT APPLICATION NUMBER: US/10/180,547
CURRENT FILING DATE: 2002-06-25
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 458
LENGTH: 734
TYPE: PRT
ORGANISM: Homo Sapien
US-10-180-547-458

Query Match 33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPFR 7
|||||
Db 140 LGPFR 144

RESULT 185
US-10-180-549-458

Sequence 458, Application US/10180549
Publication No. US20030032122A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C151
CURRENT APPLICATION NUMBER: US/10/180,549

```

; CURRENT FILING DATE: 2002-06-25
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-180-549-458

Query Match
; 33.3%; Score 5; DB 9; Length 734;
; Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPFR 7
; 11111
Db 140 LGPFR 144

RESULT 186
US-10-180-555-458
; Sequence 458, Application US/10180555
; Publication No. US20030032123A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C159
; CURRENT FILING DATE: 2002-06-25
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-180-555-458

Query Match
; 33.3%; Score 5; DB 9; Length 734;
; Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPFR 7
; 11111
Db 140 LGPFR 144

RESULT 187
US-10-180-559-458
; Sequence 458, Application US/10180559
; Publication No. US20030032124A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
```

```

; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C159
; CURRENT FILING DATE: 2002-06-25
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-180-559-458

Query Match
; 33.3%; Score 5; DB 9; Length 734;
; Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPFR 7
; 11111
Db 140 LGPFR 144

RESULT 188
US-10-181-000-458
; Sequence 458, Application US/10181000
; Publication No. US20030032125A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C17
; CURRENT FILING DATE: 2002-06-25
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-181-000-458

Query Match
; 33.3%; Score 5; DB 9; Length 734;
; Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPFR 7
; 11111
Db 140 LGPFR 144

RESULT 189
US-10-183-010-458
; Sequence 458, Application US/10183010
; Publication No. US20030032126A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
```

APPLICANT: Wood,William I.
APPLICANT: Zhang,Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C164
CURRENT APPLICATION NUMBER: US/10/183,010
CURRENT FILING DATE: 2002-06-26
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 458
LENGTH: 734
TYPE: PRT
ORGANISM: Homo Sapien
US-10-183-010-458

Query Match 33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6,4e-02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPHR 7
|||||
Db 140 LGPHR 144

RESULT 190
US-10-183-012-458
Sequence 458, Application US/10183012
Publication No. US20030032127A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Matanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3430R1C171
CURRENT APPLICATION NUMBER: US/10/183,012
CURRENT FILING DATE: 2002-06-26
PRIOR APPLICATION NUMBER: 10/052586
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063540
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063564
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063734
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063870
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31

PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066120
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066466
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066772
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069335
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069425
PRIOR FILING DATE: 1997-12-12
PRIOR APPLICATION NUMBER: 60/069870
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/068017
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/078866
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643

;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085580
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085582
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085700
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/086023
;; PRIOR FILING DATE: 1998-05-18
;; PRIOR APPLICATION NUMBER: 60/086392
;; PRIOR FILING DATE: 1998-05-22
;; PRIOR APPLICATION NUMBER: 60/086486
;; PRIOR FILING DATE: 1998-05-22
;; PRIOR APPLICATION NUMBER: 60/087098
;; PRIOR FILING DATE: 1998-05-28
;; PRIOR APPLICATION NUMBER: 60/087208
;; PRIOR FILING DATE: 1998-05-28
;; PRIOR APPLICATION NUMBER: 60/087609
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087759
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087827
;; PRIOR FILING DATE: 1998-06-03
;; PRIOR APPLICATION NUMBER: 60/088025
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088028
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088029
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088033
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088167
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088202
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088212
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088217
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088326
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088655
;; PRIOR FILING DATE: 1998-06-09
;; PRIOR APPLICATION NUMBER: 60/088722
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088738
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088740
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088811
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088824
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088825
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088826
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088861
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088863
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088876
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089090
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089105
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089512
;; PRIOR FILING DATE: 1998-06-16

;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653

Query Match 33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPFR 7
|||||
DB 140 LGPFR 144

RESULT 191
US-10-184-614-458
Sequence 458, Application US/10184614
Publication No. US20030032128A1
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jitan
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C184

CURRENT FILING DATE: 2225-06-27
Prior Application removed - See file wrapper or Palm

NUMBER OF SEQ ID NOS: 612
SEQ ID NO 458
LENGTH: 734
TYPE: PRT
ORGANISM: Homo Sapien

US-10-184-614-458

Query Match 33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPFR 7
|||||
DB 140 LGPFR 144

RESULT 192
US-10-184-623-458
Sequence 458, Application US/10184623
Publication No. US20030032129A1
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jitan
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C184

```
; FILE REFERENCE: P3430RIC210
; CURRENT APPLICATION NUMBER: US/10/184,623
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-184-623-458
```

```
Query Match          33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      3 LGPFR 7
        |||||
Db      140 LGPFR 144
```

```
RESULT 193
US-10-184-635-458
; Sequence 458, Application US/10184635
```

```
; Publication No. US20030032130A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430RIC215
; CURRENT APPLICATION NUMBER: US/10/184,635
; PRIOR FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-184-635-458
```

```
Query Match          33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      3 LGPFR 7
        |||||
Db      140 LGPFR 144
```

```
RESULT 194
US-10-184-637-458
; Sequence 458, Application US/10184637
```

```
; Publication No. US20030032131A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
```

```
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430RIC206
; CURRENT APPLICATION NUMBER: US/10/184,637
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-184-637-458
```

```
Query Match          33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      3 LGPFR 7
        |||||
Db      140 LGPFR 144
```

```
RESULT 195
US-10-184-646-458
; Sequence 458, Application US/10184646
```

```
; Publication No. US20030032132A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430RIC221
; CURRENT APPLICATION NUMBER: US/10/184,646
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-184-646-458
```

```
Query Match          33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      3 LGPFR 7
        |||||
Db      140 LGPFR 144
```

```
RESULT 196
US-10-184-647-458
; Sequence 458, Application US/10184647
```

```
; Publication No. US20030032133A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
```

```
; APPLICANT: Smith,Victoria
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C212
; CURRENT APPLICATION NUMBER: US/10/184,647
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO: 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-184-647-458

Query Match          33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPHR 7
    |||||
Db 140 LGPHR 144

RESULT 197
US-10-184-652-458
; Sequence 458, Application US/10184652
; Publication No. US20030032134A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Chen,Jian
; APPLICANT: Desnoyers,Luc
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Pan,James
; APPLICANT: Smith,Victoria
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C187
; CURRENT APPLICATION NUMBER: US/10/184,652
; CURRENT FILING DATE: 2002-06-27
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO: 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-184-652-458

Query Match          33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPHR 7
    |||||
Db 140 LGPHR 144

RESULT 198
US-10-187-594-458
; Sequence 458, Application US/10187594
; Publication No. US20030032135A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Chen,Jian
; APPLICANT: Desnoyers,Luc
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Pan,James
; APPLICANT: Smith,Victoria
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C212
; CURRENT APPLICATION NUMBER: US/10/187,594
; CURRENT FILING DATE: 2002-07-01
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO: 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-187-594-458

Query Match          33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPHR 7
    |||||
Db 140 LGPHR 144

RESULT 199
US-10-187-596-458
; Sequence 458, Application US/10187596
; Publication No. US20030032136A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Chen,Jian
; APPLICANT: Desnoyers,Luc
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Pan,James
; APPLICANT: Smith,Victoria
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C243
; CURRENT APPLICATION NUMBER: US/10/187,596
; CURRENT FILING DATE: 2002-07-02
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO: 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-187-596-458

Query Match          33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPHR 7
    |||||
Db 140 LGPHR 144

RESULT 200
US-10-187-745-458
; Sequence 458, Application US/10187745
; Publication No. US20030032137A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
```

```
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Pan,James
; APPLICANT: Smith,Victoria
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C250
; CURRENT APPLICATION NUMBER: US/10/187,594
; CURRENT FILING DATE: 2002-07-01
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO: 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-187-594-458

Query Match          33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPHR 7
    |||||
Db 140 LGPHR 144

RESULT 199
US-10-187-596-458
; Sequence 458, Application US/10187596
; Publication No. US20030032136A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Chen,Jian
; APPLICANT: Desnoyers,Luc
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Pan,James
; APPLICANT: Smith,Victoria
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C243
; CURRENT APPLICATION NUMBER: US/10/187,596
; CURRENT FILING DATE: 2002-07-02
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO: 458
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-187-596-458

Query Match          33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPHR 7
    |||||
Db 140 LGPHR 144

RESULT 200
US-10-187-745-458
; Sequence 458, Application US/10187745
; Publication No. US20030032137A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
```


APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3430R1C247
CURRENT APPLICATION NUMBER: US/10/187,745
CURRENT FILING DATE: 2002-07-02
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 458
LENGTH: 734
TYPE: PRT
ORGANISM: Homo Sapien
US-10-187-745-458

Query Match 33.3%; Score 5; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHR 7
11111
Db 140 LGPHR 144

Search completed: March 10, 2003, 14:38:47
Job time: 103 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2003, 14:23:54 ; Search time 37 Seconds
(without alignments)
38.973 Million cell updates/sec

Title: US-09-689-159a-2_COPY_346_360
Perfect score: 15
Sequence: 1 SHLGPHTSPESRAA 15

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 200 summaries

Database : PIR_73:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15	100.0	463	2	S63683 - presenilin 1-463 -
2	15	100.0	467	2	S58396 - presenilin 1, spli
3	15	100.0	467	2	I78388 - S182 protein - mou
4	8	53.3	347	2	H75253 - hypothetical prote
5	7	46.7	327	2	H87541 - hypothetical prote
6	7	46.7	343	2	E95911 - probable transcrip
7	7	46.7	463	2	JC5081 - presenilin 1 prote
8	7	46.7	467	2	JC5080 - presenilin 1 prote
9	7	46.7	512	2	S73457 - methionine-tRNA 11
10	6	40.0	87	2	F45714 - probable regulator
11	6	40.0	87	2	D61547 - hypothetical prote
12	6	40.0	138	2	D95289 - hypothetical prote
13	6	40.0	139	2	T26722 - hypothetical prote
14	6	40.0	168	2	E75257 - hypothetical prote
15	6	40.0	240	2	E45714 - molybdenum cofacto
16	6	40.0	241	2	C46181 - px-tax-ort II (alt
17	6	40.0	257	2	D96016 - probable exodeoxyr
18	6	40.0	262	2	F87498 - exodeoxyribonuclea
19	6	40.0	301	2	T37031 - hypothetical prote
20	6	40.0	364	2	A82860 - DNA replication an
21	6	40.0	382	2	S40693 - opsin rh3 - fruit
22	6	40.0	401	2	T34535 - multistage-efflux t
23	6	40.0	407	2	T34535 - conserved hypotet
24	6	40.0	429	2	D87297 - histidyl-tRNA synt
25	6	40.0	433	2	H98112 - histidine-tRNA 11g
26	6	40.0	433	2	JC5390 - presenilin-alpha 1
27	6	40.0	475	2	S31927 - drop protein - hu
28	6	40.0	617	2	T03510 - hypothetical prote
29	6	40.0	617	2	T03510 - hypothetical prote

30	6	40.0	640	2	AE1895
31	6	40.0	805	2	T64207
32	6	40.0	817	1	RRVGT
33	6	40.0	818	1	RRVGR
34	6	40.0	851	2	S52717
35	6	40.0	865	1	VGBERB
36	6	40.0	889	2	T30715
37	6	40.0	981	2	B88794
38	6	40.0	1048	2	T30815
39	6	40.0	1098	1	PFMSRB
40	6	40.0	1106	1	PFHUGB
41	6	40.0	1680	2	T01367
42	6	40.0	1711	1	A47392
43	6	40.0	2664	2	T28626
44	5	33.3	18	2	S58277
45	5	33.3	26	2	T04371
46	5	33.3	39	2	I37554
47	5	33.3	51	2	F86491
48	5	33.3	51	2	C72131
49	5	33.3	56	2	G72355
50	5	33.3	69	1	QOECF7
51	5	33.3	69	2	C90713
52	5	33.3	69	2	G85563
53	5	33.3	76	2	D72667
54	5	33.3	82	2	S28280
55	5	33.3	87	2	AD3436
56	5	33.3	102	1	CCEG
57	5	33.3	105	2	AB0318
58	5	33.3	118	2	S70089
59	5	33.3	122	2	A05114
60	5	33.3	124	2	S43840
61	5	33.3	127	2	AH3491
62	5	33.3	134	2	S23003
63	5	33.3	135	2	C82821
64	5	33.3	136	2	B84990
65	5	33.3	145	2	AD2345
66	5	33.3	150	2	T08734
67	5	33.3	157	2	A97573
68	5	33.3	160	2	A71062
69	5	33.3	163	2	E84172
70	5	33.3	165	2	D87241
71	5	33.3	165	2	T16984
72	5	33.3	166	2	I49694
73	5	33.3	166	2	T84740
74	5	33.3	168	2	S68480
75	5	33.3	173	2	F75274
76	5	33.3	173	2	S14747
77	5	33.3	176	2	T36394
78	5	33.3	183	2	A11820
79	5	33.3	185	2	T01887
80	5	33.3	189	2	G90855
81	5	33.3	189	2	A86369
82	5	33.3	191	2	A82613
83	5	33.3	191	2	T28682
84	5	33.3	197	2	B72471
85	5	33.3	198	2	A45067
86	5	33.3	201	2	A12631
87	5	33.3	201	2	B86488
88	5	33.3	206	2	T07771
89	5	33.3	206	2	T09503
90	5	33.3	206	2	AF0498
91	5	33.3	207	1	T17623
92	5	33.3	208	2	A48567
93	5	33.3	214	2	AB2157
94	5	33.3	215	2	F82627
95	5	33.3	215	2	T16393
96	5	33.3	218	2	H83475
97	5	33.3	219	2	T42605
98	5	33.3	224	2	A87094
99	5	33.3	227	2	D85018
100	5	33.3	228	2	D95865
101	5	33.3	230	2	S41043
102	5	33.3	230	2	T40233

hypothetical prote
preprotein translo
RNA-directed RNA p
RNA-directed RNA p
glycoprotein B pre
protein K07P5.12a
platelet-derived g
platelet-derived g
hypothetical prote
chromodomain-helic
variant-specific s
insulin-like growt
thymalitin-like pro
homeobox - human
hypothetical prote
hypothetical prote
hypothetical prote
yeb protein - Esc
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
cytochrome c (wall
conserved hypotet
kora protein - Amy
hypothetical prote
NADH2 dehydrogenas
hypothetical cytos
trak protein - Esc
NADH2 dehydrogenas
50S ribosomal prot
hypothetical prote
hypothetical prote
nitrogen regulator
hypothetical prote
hypothetical prote
VPS29-like phospho
transcription fact
glucokinase - mous
succinate dehydrog
conserved hypotet
sphingomyelin phos
probable pantoate-
hypothetical prote
hypothetical prote
hypothetical prote
conserved hypotet
hypothetical prote
hypothetical prote
laminin B1 chain v
hypothetical prote
hypothetical prote
probable heme tran
probable heme tran
hypothetical prote
hypothetical prote
calmodulin-ubiquit
hypothetical prote
phosphoglycerate m
hypothetical prote
probable transcrip
envelope protein -
ATP-dependent CIP
probable hypotet
probable pentose-5
rml protein - fis
Rumip - flission ye

103	5	33.3	230	2	A54501	sporulated oocyst
104	5	33.3	238	2	C72576	probable glutamine
105	5	33.3	244	2	B88115	protein F53C3.4 (l
106	5	33.3	244	2	AG2071	hypothetical prote
107	5	33.3	247	2	T33469	hypothetical prote
108	5	33.3	247	2	H69030	hypothetical prote
109	5	33.3	247	2	T17311	coenzyme PQ synth
110	5	33.3	247	2	B97669	hypothetical prote
111	5	33.3	247	2	AG2893	hypothetical prote
112	5	33.3	249	2	A70745	conserved hypotet
113	5	33.3	249	2	T16924	probable transcrip
114	5	33.3	252	2	B48725	hypothetical prote
115	5	33.3	263	2	AG2910	MDV specific prote
116	5	33.3	265	2	A87345	hypothetical prote
117	5	33.3	266	2	A12762	transcription regu
118	5	33.3	267	2	E97685	conserved hypotet
119	5	33.3	268	2	AD0898	monofunctional bio
120	5	33.3	270	2	G75411	PTS-transport faml
121	5	33.3	271	2	C96010	probable transpos
122	5	33.3	271	2	H97543	conserved hypotet
123	5	33.3	272	2	AG2215	hypothetical prote
124	5	33.3	274	2	T22993	hypothetical prote
125	5	33.3	275	1	G69963	hypothetical prote
126	5	33.3	278	1	B69025	lipoprotein SpoII
127	5	33.3	279	2	E83456	conserved hypotet
128	5	33.3	280	2	C70642	hypothetical prote
129	5	33.3	280	2	T28857	probable ribosomal
130	5	33.3	284	2	T13621	hypothetical prote
131	5	33.3	286	2	B84807	probable RNA-bind
132	5	33.3	287	2	T15779	hypothetical prote
133	5	33.3	291	2	G70605	hypothetical prote
134	5	33.3	292	2	T00996	probable hydrolase
135	5	33.3	295	2	A43633	En/spm-like transp
136	5	33.3	295	2	AF0098	host-inducible pro
137	5	33.3	297	2	T13317	hypothetical prote
138	5	33.3	299	2	F90741	hypothetical prote
139	5	33.3	299	2	A85592	hypothetical prote
140	5	33.3	299	2	AH0602	probable formate a
141	5	33.3	301	1	A54687	transcription fact
142	5	33.3	301	1	A48880	transcription fact
143	5	33.3	301	1	B84282	8-oxoguanine DNA g
144	5	33.3	302	2	T13457	hypothetical prote
145	5	33.3	302	2	F84329	hypothetical prote
146	5	33.3	302	2	S75481	polysialic acid tr
147	5	33.3	303	2	G64405	tetrahydromethanop
148	5	33.3	306	2	B70835	hypothetical prote
149	5	33.3	308	1	H64819	formate acetyltran
150	5	33.3	309	2	T41889	PE38 orf153 - Bomb
151	5	33.3	314	2	T26531	WeeP/TsgA/Cpsf fam
152	5	33.3	316	2	F87260	ATP sulfurylase, s
153	5	33.3	317	2	F82672	phospholipase A (l
154	5	33.3	320	1	AB0330	immediate-early pr
155	5	33.3	321	1	A43681	hypothetical prote
156	5	33.3	321	2	C72869	hypothetical prote
157	5	33.3	321	2	S55640	hypothetical prote
158	5	33.3	323	2	T13005	hypothetical prote
159	5	33.3	326	2	F96522	hypothetical prote
160	5	33.3	328	2	A84291	ornithine cyclodea
161	5	33.3	330	2	A40855	homeotic protein H
162	5	33.3	332	2	S44743	C02D5.1 protein -
163	5	33.3	332	2	A70388	hydrogenase expres
164	5	33.3	332	2	I57032	gene Tlx-1 protein
165	5	33.3	332	2	AG0809	probable membrane
166	5	33.3	334	2	C22735	hypothetical nox2
167	5	33.3	334	2	AB3342	conserved hypotet
168	5	33.3	336	2	B86429	F26G16.5 protein -
169	5	33.3	336	2	T44988	oxidoreductase (im
170	5	33.3	338	2	E83059	ketol acid reducto
171	5	33.3	338	2	A83835	NAOH oxidase BH18
172	5	33.3	338	2	C98242	hypothetical prote
173	5	33.3	338	2	AI13043	oxidoreductase Atu
174	5	33.3	342	2	B85757	partial probable o
175	5	33.3	342	2	AB3116	transcription regu

ALIGNMENTS

bifunctional short
hypothetical prote
hypothetical prote
hypothetical prote
unknown protein 11
probable UDP-3-O-l
hypothetical prote
hypothetical prote
prostaglandin E2 r
EP2 prostaglandin
hypothetical prote
homeotic protein D
acriflavine resista
CD44 membrane glyc
actin modulator pr
probable FMN oxido
probable zinc-bind
hypothetical prote
probable bacteriop
hypothetical prote
succinate dehydrog
hypothetical prote
chitinase (Ec 3.2.
hypothetical prote
transforming prote

RESULT 1
S63683
presenilin I-463 - human
C:Species: Homo sapiens (man)
C>Date: 20-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 29-Sep-1999
C:Accession: S63683
R:Sahara, N.; Yabagi, Y.; Takagi, H.; Kondo, T.; Okochi, M.; Usami, M.; Shirasawa, T.
FEBS Lett. 381, 7-11, 1996
A>Title: Identification and characterization of presenilin I-467, I-463 and I-374.
A:Reference number: S63683; MUID:96193901; PMID:8641442
A:Accession: S63683
A>Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-463 <SAB>
A:Cross-references: EMBL:U040379; NID:g1244637; PIDN:AAB05894.1; PID:g1244638
C:Superfamily: presenilin

Query Match 100.0%; Score 15; DB 2; Length 463;
Best Local Similarity 100.0%; Pred. No. 2.7e-09;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPHRSTPESRAA 15
DB 342 SHLGPHRSTPESRAA 356

RESULT 2

S58396
presenilin 1, splice form 467 - human
N:Alternative names: Alzheimer's disease protein 3; protein S182
C:Species: Homo sapiens (man)
C>Date: 29-Jan-1998 #sequence_revision 13-Feb-1998 #text_change 29-Sep-1999
C:Accession: S58396; S71401; S71402
R:Sherrington, R.; Rogeev, E.I.; Liang, Y.; Rogeeva, E.A.; Levesque, G.; Ikeda, M.; C
ero, I.; Plinesil, L.; Nee, L.; Chinnakoy, I.; Pollen, D.; Brookes, A.; Sansau, P.; Po
E.; Rommens, J.M.; St George-Hyslop, P.H.
Nature 375, 754-760, 1995
A>Title: Cloning of a gene bearing missense mutations in early-onset familial Alzheimer
A:Reference number: I58095; MUID:95319502; PMID:7596406
A:Accession: S58396
A:Molecule type: mRNA
A:Residues: 1-467 <SHE>
A:Cross-references: EMBL:L42110; NID:g904118; PIDN:AAB46416.1; PID:g904119

R; Vidal, R.; Ghiso, J.; Wisniewski, T.; Frangione, B.
 FEBS Lett. 393, 19-23, 1996
 A:Title: Alzheimer's presenilin 1 gene expression in platelets and megakaryocytes. Ident
 A:Reference number: S71401, MUID:96397521, PMID:8804415
 A:Accession: S71401
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 24-32;254-256,290-292;316-317,376-379 <VIM>
 A:Experimental source: Damt megakaryocytic cell line (ATCC CRL-9792) and platelets
 C:Genetics:
 A:Gene: GDB:PSEN1, AD3, FAD: S182, PS1
 A:Cross-references: GDB:135682; OMIM:104311
 A:Map position: 14q24.3-14q24.3
 C:Superfamily: presenilin
 C:Keywords: alternative splicing; Alzheimer's disease; glycoprotein; transmembrane prote
 F:82-100/Domain: transmembrane #status predicted <TM1>
 F:133-154/Domain: transmembrane #status predicted <TM2>
 F:164-185/Domain: transmembrane #status predicted <TM3>
 F:195-213/Domain: transmembrane #status predicted <TM4>
 F:221-238/Domain: transmembrane #status predicted <TM5>
 F:244-264/Domain: transmembrane #status predicted <TM6>
 F:281-301/Domain: transmembrane #status predicted <TM7>
 F:408-428/Domain: transmembrane #status predicted <TM8>
 F:433-453/Domain: transmembrane #status predicted <TM9>
 F:279,405/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 15; DB 2; Length 467;
 Best Local Similarity 100.0%; Pred. No. 2.7e-09;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLPHRSTPESRAA 15
 |||||
 Db 346 SHLPHRSTPESRAA 360

RESULT 3
 S182 protein - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 29-Sep-1999
 C:Accession: I78388
 R:Sherrington, R.; Rogaev, E.I.; Liang, Y.; Rogaeva, E.A.; Levesque, G.; Ikeda, M.; Chi,
 ero, I.; Pinessi, L.; Nee, L.; Chumakov, I.; Pollen, D.; Brookes, A.; Sanseau, P.; Polin
 E.; Rommens, J.M.; St George-Hyslop, P.H.
 Nature 375, 754-760, 1995
 A:Title: Cloning of a gene bearing missense mutations in early-onset familial Alzheimer
 A:Reference number: I58095; MUID:95319502; PMID:756406
 A:Accession: I78388
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-467 <RES>
 A:Cross-references: GB:I42177; NID:9904129; PIDN:AAC42094.1; PID:9904130
 C:Superfamily: presenilin

Query Match 100.0%; Score 15; DB 2; Length 467;
 Best Local Similarity 100.0%; Pred. No. 2.7e-09;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLPHRSTPESRAA 15
 |||||
 Db 346 SHLPHRSTPESRAA 360

RESULT 4
 H75253
 hypothetical protein - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
 C:Accession: H75253
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 M., Shen, M.; Yamathayan, J.J.; Lam, P.; McDonald, L.; Uitterback, T.; Zaleski, C.; Ma
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266
 A:Accession: H75253
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-347 <WHI>
 A:Cross-references: GB:AE002089; GB:AE000513; NID:96460427; PIDN:AAF12144.1; PID:964
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DR2600
 A:Map position: 1

Query Match 53.3%; Score 8; DB 2; Length 347;
 Best Local Similarity 100.0%; Pred. No. 0.14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPESRA 14
 |||||
 Db 274 RSTPESRA 281

RESULT 5
 F87541
 hypothetical protein CC2359 [Imported] - Caulobacter crescentus
 C:Species: Caulobacter crescentus
 C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 17-May-2002
 C:Accession: F87541
 R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg
 B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Hatt, D.H.; K
 u, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A:Title: Complete Genome Sequence of Caulobacter crescentus.
 A:Reference number: A87249; MUID:21173698; PMID:11259647
 A:Accession: F87541
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-327 <STO>
 A:Cross-references: GB:AE005673; NID:913423888; PIDN:AAK2430.1; GSPDB:GN00148
 C:Genetics:
 A:Gene: CC2359
 C:Superfamily: N4-(beta-N-acetylglucosaminyl)-L-asparaginase

Query Match 46.7%; Score 7; DB 2; Length 327;
 Best Local Similarity 100.0%; Pred. No. 1.7;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRAA 15
 |||||
 Db 139 TPESRAA 145

RESULT 6
 E95911
 probable transcription regulator, LacI family protein [Imported] - Sinorhizobium meli
 C:Species: Sinorhizobium meli
 C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
 C:Accession: E95911
 R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; He
 Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
 A:Title: The complete sequence of the 1.683-kb pSymb megaplasmid from the N2-fixing
 A:Reference number: A95842; MUID:21396508; PMID:11481431
 A:Accession: E95911
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-343 <KUR>
 A:Cross-references: GB:AL591985; PIDN:CAC48957.1; PID:915140442; GSPDB:GN00167
 A:Experimental source: strain 1021, megaplasmid pSymb
 R:Gilbert, F.; Finan, T.M.; Long, S.R.; Pulver, A.; Abola, P.; Ampe, F.; Barloy-Hub
 pel, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R
 L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lela
 hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,

A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 A:Contents: annotation
 C:Genetics:
 A:Gene: Smb20817
 A:Genome: plasmid

Query Match 46.7%; Score 7; DB 2; Length 343;
 Best Local Similarity 100.0%; Pred. No. 1.8;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 9 TPESRAA 15
 Db 285 TPESRAA 291

RESULT 7
 JCS081
 presentin 1 protein isoform 463 - lesser mouse lemur
 C:Species: Microcebus murinus (lesser mouse lemur)
 C:Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 13-Sep-1998
 C:Accession: JCS081
 R:Calenda, A.; Mestre-Frances, N.; Czech, C.; Pradier, L.; Pelter, A.; Bons, N.; Bellis, Biochem. Biophys. Res. Commun. 228, 430-439, 1996
 A:Title: Molecular cloning, sequencing and brain expression of the presentin 1 gene in A:Reference number: JCS080; MUID:97079199; PMID:8920931
 A:Contents: brain
 A:Accession: JCS081
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-463 <CAL>
 A:Cross-references: EMBL:Z71333
 C:Comment: This protein is an intermembrane protein with seven transmembrane domains. It C:Genetics:
 A:Gene: psi
 A:Map position: 14
 C:Superfamily: presentin
 C:Keywords: transmembrane protein
 F:78-96/Domain: transmembrane #status predicted <TM1>
 F:129-150/Domain: transmembrane #status predicted <TM2>
 F:160-181/Domain: transmembrane #status predicted <TM3>
 F:191-209/Domain: transmembrane #status predicted <TM4>
 F:217-234/Domain: transmembrane #status predicted <TM5>
 F:240-257/Domain: transmembrane #status predicted <TM6>
 F:404-424/Domain: transmembrane #status predicted <TM7>

Query Match 46.7%; Score 7; DB 2; Length 463;
 Best Local Similarity 100.0%; Pred. No. 2.4;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 LGPHRST 9
 Db 344 LGPHRST 350

RESULT 8
 JCS080
 presentin 1 protein isoform 467 - lesser mouse lemur
 C:Species: Microcebus murinus (lesser mouse lemur)
 C:Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 20-Jun-2000
 C:Accession: JCS080
 R:Calenda, A.; Mestre-Frances, N.; Czech, C.; Pradier, L.; Pelter, A.; Bons, N.; Bellis, Biochem. Biophys. Res. Commun. 228, 430-439, 1996
 A:Title: Molecular cloning, sequencing, and brain expression of the presentin 1 gene in A:Reference number: JCS080; MUID:97079199; PMID:8920931
 A:Accession: JCS080
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-467 <CAL>
 A:Cross-references: EMBL:Z71333; NID:g1707591; PIDN:CA95930.1; PID:g1707592
 A:Experimental source: brain
 C:Comment: This protein is an intergal membrane protein with seven transmembrane domain C:Genetics:

A:Gene: psi
 A:Map position: 14
 C:Superfamily: presentin
 C:Keywords: transmembrane protein
 F:82-100/Domain: transmembrane #status predicted <TM1>
 F:133-154/Domain: transmembrane #status predicted <TM2>
 F:164-185/Domain: transmembrane #status predicted <TM3>
 F:195-213/Domain: transmembrane #status predicted <TM4>
 F:221-238/Domain: transmembrane #status predicted <TM5>
 F:244-261/Domain: transmembrane #status predicted <TM6>
 F:408-428/Domain: transmembrane #status predicted <TM7>

Query Match 46.7%; Score 7; DB 2; Length 467;
 Best Local Similarity 100.0%; Pred. No. 2.4;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 LGPHRST 9
 Db 348 LGPHRST 354

RESULT 9
 S73457
 methionine-tRNA ligase (EC 6.1.1.10) mets - Mycoplasma pneumoniae (strain ATCC 29342)
 N:Alternate names: hypothetical protein B01_orf512; methionyl-tRNA synthetase mets
 C:Species: Mycoplasma pneumoniae
 A:Variety: ATCC 29342
 C:Date: 26-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 03-Jun-2002
 C:Accession: S73457
 R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkl, E.; Li, B.C.; Herrmann, R. Nucleic Acids Res. 24, 4420-4449, 1996
 A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumonia A:Reference number: S73327; MUID:97105885; PMID:8948633
 A:Accession: S73457
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-512 <HIM>
 A:Cross-references: EMBL:AE000015; GB:U00089; NID:g1673779; PIDN:AAB95779.1; PID:g167 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996 C:Genetics:
 A:Gene: mets
 A:Gene code: SGC3
 C:Superfamily: methionine-tRNA ligase
 C:Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 46.7%; Score 7; DB 2; Length 512;
 Best Local Similarity 100.0%; Pred. No. 2.6;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 9 TPESRAA 15
 Db 180 TPESRAA 186

RESULT 10
 F45714
 probable regulatory protein p13 II, alternative splice form - human T-cell lymphotropic C:Species: human T-cell lymphotropic virus type 1, HTLV-1
 C:Date: 22-Sep-1993 #sequence_revision 25-Apr-1997 #text_change 23-May-1997
 C:Accession: F45714
 R:Gesash, A.; Boerl, E.; Yanagihara, R.; Gallo, R.C.; Franchini, G. J. Virol. 67, 1015-1023, 1993
 A:Title: Complete nucleotide sequence of a highly divergent human T-cell leukemia (1y from other geographical regions
 A:Reference number: A45714; MUID:93124536; PMID:8419636
 A:Contents: HTLV-IMEL5
 A:Accession: F45714
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: nucleic acid
 A:Residues: 1-87 <GES>
 A:Note: sequence extracted from NCBI backbone (NCBI:P:122473)
 Query Match 40.0%; Score 6; DB 2; Length 87;

Best Local Similarity 100.0%; Pred. No. 7;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HLGPHR 7
|||||
Db 58 HLGPHR 63

RESULT 11

hypothetical protein II (px region) - human T-cell lymphotropic virus type 1 (isolate HA
C:Species: human T-cell lymphotropic virus type 1, HTLV-1
C:Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 08-Oct-1999
R:Tsujimoto, A.; Teruuchi, T.; Imanura, J.; Shimotohno, K.; Miyoshi, I.; Miwa, M.
Mol. Biol. Med. 5, 29-42, 1988
A:Title: Nucleotide sequence analysis of a provirus derived from HTLV-1-associated myelo
A:Reference number: A61547; MUID:88232270; PMID:2897612
A:Accession: D61547
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-87 <TSU>
A:Cross-references: GB:M37301; NID:9541634; PIDN:AAA5391.1; PID:9541638

Query Match 40.0%; Score 6; DB 2; Length 87;
Best Local Similarity 100.0%; Pred. No. 7;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HLGPHR 7
|||||
Db 58 HLGPHR 63

RESULT 12

hypothetical protein Sma0412 [Imported] - Sinorhizobium meliloti (strain 1021) magaplast
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: D95289
R:Barrett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Stutzky, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilot
A:Reference number: A95262; MUID:21396509; PMID:11481432
A:Accession: D95289
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-138 <KUR>
A:Cross-references: GB:AE006469; PIDN:AAK64878.1; PID:G14523295; GSPDB:GN00165

R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenberg, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: Sma0412
A:Genome: plasmid

Query Match 40.0%; Score 6; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHRS 8
|||||
Db 109 LGPHRS 114

RESULT 13

T26722

hypothetical protein Y39A1A.18 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T26722
R:Wall, M.
submitted to the EMBL Data Library, September 1998
A:Reference number: Z20257
A:Accession: T26722
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-139 <MI>
A:Cross-references: EMBL:AL031633; PIDN:CAA21011.1; GSPDB:GN00021; CESP:Y39A1A.18
A:Experimental source: clone Y39A1A
C:Genetics:
A:Gene: CESP:Y39A1A.18
A:Map position: 3
A:Introns: 37/3; 79/3

Query Match 40.0%; Score 6; DB 2; Length 139;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRA 14
|||||
Db 19 TPESRA 24

RESULT 14

E75257
molybdenum cofactor biosynthesis protein C - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: E75257
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J
; M.; Shen, H.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: E75257
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-168 <MHI>
A:Cross-references: GB:AE002086; GB:AE000513; NID:96460395; PIDN:AAF12111.1; PID:9646
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR2571
A:Map position: 1
C:Superfamily: molybdenum cofactor biosynthesis protein C

Query Match 40.0%; Score 6; DB 2; Length 168;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRAA 15
|||||
Db 44 PESRAA 49

RESULT 15

EA5714

probable regulatory function protein p30 II, alternative splice form - human T-cell 1
C:Species: human T-cell lymphotropic virus type 1, HTLV-1
C:Date: 22-Sep-1993 #sequence_revision 25-Apr-1997 #text_change 23-May-1997
C:Accession: EA5714
R:Gessein, A.; Boeri, E.; Yanagihara, R.; Gallo, R.C.; Franchini, G.
J. Virol. 67, 1015-1023, 1993

A:Title: Complete nucleotide sequence of a highly divergent human T-cell leukemia (Ty
from other geographical regions.
A:Reference number: A45714; MUID:93124536; PMID:8419636
A:Contents: HTLV-IMEL5
A:Accession: EA5714

A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-240 <GEN>
A:Note: sequence extracted from NCBI backbone (NCBI:122472)

Query Match 40.0%; Score 6; DB 2; Length 240;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HLGPRH 7
|||||
Db 211 HLGPRH 216

RESULT 16
C46181
px-tax-orf II (alternatively spliced) - human T-cell lymphotropic virus type 1
C:Species: human T-cell lymphotropic virus type 1, HTLV-1
C:Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 17-Nov-2000
C:Accession: C46181; D46181
R:Koralnik, I.J.; Gessalin, A.; Klotman, M.E.; Lo Monaco, A.; Berneman, Z.N.; Franchini, P.
Proc. Natl. Acad. Sci. U.S.A. 89, 8813-8817, 1992
A:Title: Protein isoforms encoded by the pX region of human T-cell leukemia/lymphotropic
A:Reference number: A46181; MUID:92409607; PMID:1528897
A:Accession: C46181
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-241 <KOR>
A:Note: sequence extracted from NCBI backbone (NCBI:114307)

A:Accession: D46181
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 155-241 <K02>
A:Note: sequence extracted from NCBI backbone (NCBI:114308)

Query Match 40.0%; Score 6; DB 2; Length 241;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HLGPRH 7
|||||
Db 212 HLGPRH 217

RESULT 17
D96016
probable exodeoxyribonuclease III (EC 3.1.11.2) [imported] - Sinorhizobium meliloti (str
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 14-Sep-2001
C:Accession: D96016
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Hernat
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: D96016
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-257 <KOR>
A:Cross-references: GB:AL591985; PIDN:CAC49796.1; PID:915141283; GSPDB:GN00167
A:Experimental source: strain 1021, megaplasmid pSymB
R:Galibert, F.; Finan, T.M.; Long, S.R.; Punter, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowle, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yen, K.
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: xthA4; SMD20689
C:Superfamily: exodeoxyribonuclease III

C:Keywords: hydrolase

Query Match 40.0%; Score 6; DB 2; Length 257;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRAA 15
|||||
Db 171 PESRAA 176

RESULT 18
F87498
exodeoxyribonuclease III [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C:Accession: F87498
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: F87498
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-262 <STO>
A:Cross-references: GB:AE05673; NID:913423482; PIDN:AKK23986.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC2011
C:Superfamily: exodeoxyribonuclease III

Query Match 40.0%; Score 6; DB 2; Length 262;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRAA 15
|||||
Db 176 PESRAA 181

RESULT 19
T37031
hypothetical protein SCU12.12c - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T37031
R:Murphy, D.; Harris, D.; Thomson, N.R.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A
submitted to the EMBL Data Library, August 1999
A:Reference number: Z21619
A:Accession: T37031
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-301 <MGR>
A:Cross-references: EMBL:AL109989; PIDN:CAB53424.1; GSPDB:GN00070; SCOEDB:SCU12.12c
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SCU12.12c

Query Match 40.0%; Score 6; DB 2; Length 301;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRAA 15
|||||
Db 13 PESRAA 18

RESULT 20
A82860
DNA replication and repair RecF protein XF0003 [imported] - Xylella fastidiosa (strat
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 17-Nov-2000

C:Accession: A82860
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: A82860
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-364 <SIM>
A:Cross-references: GB:AE003855; GB:AE003849; NID:99104760; PIDN:AAF82816.1; GSPDB:GN001
A:Experimental source: strain 9a5c
R:Simpton, A.J.G.; Relnach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
B:Ramos, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carer, H
as-Nero, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Lalga
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, B
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XFR0003
C:superfamily: recF protein
Query Match 40.0%; Score 6; DB 2; Length 364;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 LGPHRS 8
DB 254 LGPHRS 259
RESULT 21
S40693
opsin rh3 - fruit fly (Drosophila pseudoobscura)
C:Species: Drosophila pseudoobscura
C:Date: 19-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 13-Aug-1999
C:Accession: S40693; S65548; S24607
R:Carulli, J.P.; Hartl, D.L.
Genetics 132, 193-204, 1992
A:Title: Variable rates of evolution among Drosophila opsin genes.
A:Reference number: S40691; MUID:93012921; PMID:1398053
A:Accession: S40693
A:Molecule type: DNA
A:Residues: 1-382 <CAR>
A:Cross-references: EMBL:X65879
A:Note: the authors translated the codon GCA for residue 25 as Gly
R:Carulli, J.P.
submitted to the EMBL Data Library, November 1992
A:Reference number: S65546
A:Accession: S65548
A:Molecule type: DNA
A:Residues: 1-123, 'V', 125-241, 'Q', 243-254, 'R', 256-382 <CAMP>
A:Cross-references: EMBL:X65879; NID:99081; PIDN:CAA46710.1; PID:99082
C:Genetics:
A:Gene: FlyBase:DPse/Rh3
A:Cross-references: FlyBase:FBgn0012709
C:superfamily: vertebrate rhodopsin
Query Match 40.0%; Score 6; DB 2; Length 382;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 10 PESRAA 15
DB 362 PESRAA 367

RESULT 22
F69834
multidrug-efflux transporter homolog yhjO - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C:Accession: F69834
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
C.; Bron, S.; Bouilliet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A.; Ehrlich, S.D.; Emerson, P.T.; Ertan, K.D.; Errington, J.; Fabbre, C.; Ferrari,
Nature 399, 249-256, 1997
A:Authors: Fowler, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal
leho, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, A
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kuita, K.; Lapidus, A.; Lardinc
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
Y, M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portele
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanl
A:Authors: Schleicher, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchlye
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
A:Authors: Yoshikawa, H.F.; Zumslein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: F69834
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1401 <RUN>
A:Cross-references: GB:299109; GB:AL009126; NID:92633260; PIDN:CAB12898.1; PID:92633
A:Experimental source: strain 168
C:Genetics:
A:Gene: yhjO
C:superfamily: tetracycline resistance protein
Query Match 40.0%; Score 6; DB 2; Length 401;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 8 STPEPS 13
DB 131 STPEPS 136
RESULT 23
T34535
hypothetical protein DKFZp434H0717.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T34535
R:Blöcker, H.; Boecker, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wilmann, S.
submitted to the Protein Sequence Database, October 1999
A:Reference number: Z21539
A:Accession: T34535
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-407 <BLO>
A:Cross-references: EMBL:AL122102
A:Experimental source: adult testis; clone DKFZp434H0717
C:Genetics:
A:Note: DKFZp434H0717.1
Query Match 40.0%; Score 6; DB 2; Length 407;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 RSTPEPS 12
DB 23 RSTPEPS 28
RESULT 24
D87297
conserved hypothetical protein CC0389 [Imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: D87297
B: Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
D., J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: D87297
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-408 <STO>
A:Cross-references: GB:AE005673; NID:g13421548; PIDN:AAK22376.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC0389

Query Match 40.0%; Score 6; DB 2; Length 408;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRAA 15
|||||
DB 117 PESRAA 122

RESULT 25
C95248
histidyl-tRNA synthetase [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
C:Accession: C95248
R:Teitelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfel,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A:Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: C95248
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-429 <KUR>
A:Cross-references: GB:AE005672; PIDN:AAK76180.1; PID:g14973634; GSPDB:GN00164; TIGR:SPA
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP2121
C:Superfamily: histidine-tRNA ligase; histidine-tRNA ligase homology

Query Match 40.0%; Score 6; DB 2; Length 429;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRAA 15
|||||
DB 171 PESRAA 176

RESULT 26
H98112
histidine-tRNA ligase (EC 6.1.1.21) [imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 03-Jun-2002
C:Accession: H98112
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: H98112
A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-429 <KUR>
A:Cross-references: GB:AE007317; PIDN:AAU00733.1; PID:g15459628; GSPDB:GN00174
C:Genetics:
A:Gene: hisS
C:Superfamily: histidine-tRNA ligase; histidine-tRNA ligase homology
C:Keywords: ligase

Query Match 40.0%; Score 6; DB 2; Length 429;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 PESRAA 15
|||||
DB 171 PESRAA 176

RESULT 27
JC5390
presentin-alpha - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 04-Jun-1997 #sequence_revision 18-Jul-1997 #text_change 20-Jun-2000
C:Accession: JC5390
R:Tsujimura, A.; Yasojima, K.; Hashimoto-Gotoh, T.
Biochem. Biophys. Res. Commun. 231, 392-396, 1997
A:Title: Cloning of Xenopus presentin-alpha and -beta cDNAs and their differential e
A:Reference number: JC5390; MUID:97223465; PMID:9070286
A:Accession: JC5390
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-433 <TSU>
A:Cross-references: DDBJ:D84427; NID:g1944353; PIDN:BAU19570.1; PID:g1944354
A:Experimental source: brain
C:Comment: This protein plays a role in negative regulation of apoptotic cascades dur
C:Superfamily: presentin
F:48-66/Domain: transmembrane #status predicted <TM1>
F:99-119/Domain: transmembrane #status predicted <TM2>
F:130-149/Domain: transmembrane #status predicted <TM3>
F:161-178/Domain: transmembrane #status predicted <TM4>
F:187-203/Domain: transmembrane #status predicted <TM5>
F:210-227/Domain: transmembrane #status predicted <TM6>
F:374-394/Domain: transmembrane #status predicted <TM7>

Query Match 40.0%; Score 6; DB 2; Length 433;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STPEER 13
|||||
DB 319 STPEER 324

RESULT 28
S31927
drop3 protein - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 10-Sep-1997
C:Accession: S31927
R:Kirsch, K.
submitted to the EMBL Data Library, February 1993
A:Reference number: S31927
A:Accession: S31927
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-475 <KTR>
A:Cross-references: EMBL:X70991; NID:g38459; PID:g38460

Query Match 40.0%; Score 6; DB 2; Length 475;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPEE 12
|||||

DB 139 RSTPES 144

RESULT 29

T03510

hypothetical protein - Rhodobacter capsulatus

C:Species: Rhodobacter capsulatus

C>Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 08-Oct-1999

C:Accession: T03510

R:Vleck, C.; Paces, V.; Maltsev, N.; Paces, J.; Haselkorn, R.; Fonstein, M.

Proc. Natl. Acad. Sci. U.S.A. 94, 9384-9388, 1997

A:Title: Sequence of a 189-kb segment of the chromosome of Rhodobacter capsulatus SB1003

A:Reference number: z14955; MUID:97404404; PMID:9256491

A:Accession: T03510

A>Status: preliminary; translated from GB/EMBL/DDJ

A:Molecule type: DNA

A:Residues: 1-617 <YLC>

A:Cross-references: EMBL:AF010496; NID:93128256; PIDN:AC16163.1; PID:93128311

C:Genetics:

A:Map position: 1

Query Match

Best Local Similarity 40.0%; Score 6; DB 2; Length 617;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRAA 15

DB 470 PESRAA 475

RESULT 30

AE1895

hypothetical protein all0711 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp.

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002

C:Accession: AE1895

R:Kanehko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Matnabe, A.; Iriguchi,

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A:Title: Complete Genome Sequence of the Filamentous Nitrogen-fixing Cyanobacterium And

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AE1895

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-640 <KUR>

A:Cross-references: GB:BA000019; PIDN:BA072668.1; PID:q17130056; GSPDB:GNO0179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: all0711

Query Match

Best Local Similarity 40.0%; Score 6; DB 2; Length 640;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRAA 15

DB 142 PESRAA 147

RESULT 31

I64207

preprotein translocase secA - Mycoplasma genitalium

C:Species: Mycoplasma genitalium

C>Date: 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change 17-Nov-2000

C:Accession: I64207

R:Prasert, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.;

M.; Fuhmann, J.; Nguyen, D.; Uitterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.;

C.A.; Venter, J.C.

Science 270, 397-403, 1995

A:Title: The minimal gene complement of Mycoplasma genitalium.

A:Reference number: A64200; MUID:96026346; PMID:7569993

A:Accession: I64207

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-806 <TIGR>

A:Cross-references: GB:039687; GB:I43967; NID:93844663; PIDN:AA071290.1; PID:91045748

A:Experimental source: strain G-37

C:Comment: The "nucleotide-binding motif B" and "DEAD motif" features as annotated at

is adjacent to the identified motif and a third conserved motif is approximately 120-1

C:Genetics:

A:Genetic code: SGC3

C:Superfamily: preprotein translocase secA

C:Keywords: ATP; membrane-associated complex; P-loop; protein transport

F:102-109/Region: nucleotide-binding motif A (P-loop) #status atypical

F:205-210/Region: nucleotide-binding motif B

F:209-212/Region: DEAD motif

Query Match

Best Local Similarity 40.0%; Score 6; DB 2; Length 806;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 PESRAA 15

DB 255 PESRAA 260

RESULT 32

RVVGR

RNA-directed RNA polymerase (EC 2.7.7.48) - tomato bushy stunt virus (strain cherry)

N:Alternate names: RNA nucleotidyltransferase (RNA-directed); RNA replicase

N:Contains: 33k protein

C:Species: tomato bushy stunt virus, TBSV

C>Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 25-Oct-1996

C:Accession: A35315

R:Hearne, P.O.; Knorr, D.A.; Hillman, B.I.; Morris, T.J.

Virology 177, 141-151, 1990

A:Title: The complete genome structure and synthesis of infectious RNA from clones of

A:Reference number: A35315; MUID:90281577; PMID:2353450

A:Accession: A35315

A:Molecule type: genomic RNA

A:Residues: 1-817 <HEAR>

A:Cross-references: EMBL:M31019

A:Note: readthrough of the terminator UAG occurs between codons AAA for 296-Lys and

C:Superfamily: carnation mottle virus RNA-directed RNA polymerase; barley yellow dwarf

C:Keywords: nucleotidyltransferase; RNA biosynthesis; RNA replication

F:1-296/Product: 33k protein #status predicted <P3>

F:527-694/Domain: barley yellow dwarf virus RNA-directed RNA polymerase homology <BYL

Query Match

Best Local Similarity 40.0%; Score 6; DB 1; Length 817;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRA 14

DB 766 TPESRA 771

RESULT 33

RVVGR

RNA-directed RNA polymerase (EC 2.7.7.48) - Cymbidium ringspot virus

N:Alternate names: RNA nucleotidyltransferase (RNA-directed); RNA replicase

N:Contains: 33k protein

C:Species: Cymbidium ringspot virus

C>Date: 31-Mar-1990 #sequence_revision 01-Dec-2000 #text_change 01-Dec-2000

C:Accession: S05456; JS0268

R:Griceo, F.; Burgyn, J.; Russo, M.

Nucleic Acids Res. 17, 6383, 1989

A:Title: The nucleotide sequence of Cymbidium ringspot virus RNA.

A:Reference number: JS0268; MUID:89366663; PMID:2771646

A:Accession: S05456

A>Status: preliminary; translation not shown

A:Molecule type: genomic RNA

A:Residues: 1-818 <GRI>

A:Cross-references: EMBL:X15511; NID:959020; PIDN:CAB38439.1; PID:94469160

A:Accession: JS0268

A.Molecule type: genomic RNA
A:Residues: 1-296,298-483,'1',485-818 <GR2>
A:Cross-references: GB:X15511
A>Note: readthrough of the terminator TAG occurs between codons AAA for 296-Tys and GGA
C:Superfamily: carnation mottle virus RNA-directed RNA polymerase; barley yellow dwarf
C:Keywords: nucleotidyltransferase; RNA biosynthesis; RNA replication
F:1-296/Product: 33k protein #status predicted <TTP>
F:528-695/Domain: barley yellow dwarf virus RNA-directed RNA polymerase homology <BYD>

Query Match
Best Local Similarity 100.0%; Score 6; DB 1; Length 818;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRA 14
|||||
DB 767 TPESRA 772

RESULT 34
S52717
RNA-directed RNA polymerase (EC 2.7.7.48) - carnation Italian ringspot virus
N:Alternate names: RNA replicase
C:Species: carnation Italian ringspot virus
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 18-Jun-1999
C:Accession: S52717
R:Rubino, L.; Burgyan, J.; Russo, M.
submitted to the EMBL Data Library, March 1995
A:Description: Molecular cloning and complete nucleotide sequence of carnation Italian
A:Reference number: S52717
A:Accession: S52717
A.Molecule type: genomic RNA
A:Residues: 1-851 <RUB>
A:Cross-references: EMBL:X85215; NID:g755716; PIDN:CAA59478.1; PID:g755718
C:Superfamily: carnation mottle virus RNA-directed RNA polymerase; barley yellow dwarf
C:Keywords: nucleotidyltransferase; RNA biosynthesis; RNA replication
F:561-728/Domain: barley yellow dwarf virus RNA-directed RNA polymerase homology <BYD>

Query Match
Best Local Similarity 100.0%; Score 6; DB 2; Length 851;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRA 14
|||||
DB 800 TPESRA 805

RESULT 35
VGBERR
glycoprotein B precursor - Marek's disease virus (strain RB1B)
C:Species: Marek's disease virus
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Jun-2000
C:Accession: A32402; B32402
R:Ross, L.J.N.; Sanderson, M.; Scott, S.D.; Blms, M.M.; Doel, T.; Milne, B.
J. Gen. Virol. 70, 1789-1804, 1989
A:Title: Nucleotide sequence and characterization of the Marek's disease virus homologue
A:Reference number: A32402; MUID:89293086; PMID:2544666
A:Accession: A32402
A.Molecule type: DNA
A:Residues: 1-865 <ROS>
A:Cross-references: GB:D13713; EMBL:D00506; NID:g221836; PIDN:BA02866.1; PID:g221837
A:Accession: B32402
A.Molecule type: protein
A:Residues: 250-271;304-330 <ROS2>
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-21/Domain: signal sequence #status predicted <STG>
F:22-865/Product: glycoprotein B #status predicted <GPB>
F:709-728/Domain: transmembrane #status predicted <TN1>
F:733-752/Domain: transmembrane #status predicted <TN2>
F:27,184,332,406,425,631/Binding site: carbohydrate (asn) (covalent) #status predict

Query Match
Best Local Similarity 100.0%; Score 6; DB 1; Length 865;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRA 14
|||||
DB 174 TPESRA 179

RESULT 36
T30715
probable major core protein 113L - Molluscum contagiosum virus 1
N:Alternate names: MC113L
C:Species: Molluscum contagiosum virus 1
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 21-Jul-2000
C:Accession: T30715
R:Senkevich, T.G.; Buger, J.J.; Sisler, J.R.; Koonin, E.V.; Darai, G.; Moss, B.
Science 273, 813-816, 1996
A:Title: Genome sequence of a human tumorigenic poxvirus: Prediction of specific host
A:Reference number: Z20876; MUID:96323459; PMID:8670425
A:Accession: T30715
A:Status: Preliminary; translated from GB/EMBL/DBJ
A.Molecule type: DNA
A:Residues: 1-889 <SEN>
A:Cross-references: EMBL:U60315; NID:g1491943; PIDN:AAC55241.1; PID:g1492056
C:Genetics:
A:Note: MC113L
C:Superfamily: vaccinia virus major core protein P4a

Query Match
Best Local Similarity 100.0%; Score 6; DB 2; Length 889;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRA 14
|||||
DB 211 TPESRA 216

RESULT 37
B88794
protein K07P5.12a [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: B88794
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating bio
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C-
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;
A:Accession: B88794
A:Status: Preliminary
A.Molecule type: DNA
A:Residues: 1-981 <STO>
A:Cross-references: GB:chr_IV; PIDN:CAA94287.1; PID:g4008375; GSPDB:GN00022; CESP:K07
C:Genetics:
A:Gene: K07P5.12a
A:Map position: 4

Query Match
Best Local Similarity 100.0%; Score 6; DB 2; Length 981;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 PHRSTP 10
|||||
DB 533 PHRSTP 538

RESULT 38
T30815
platelet-derived growth factor receptor beta - Japanese pufferfish
C:Species: Fugu rubripes (Japanese pufferfish)
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 02-Jun-2000
R:How, G.F.; Venkatesh, B.; Brenner, S.

Genome Res. 6:1185-1191, 1996

A:Title: Conserved linkage between the pufferfish (*Fugu rubripes*) and human genes for platelet-derived growth factor receptor beta precursor - mouse

A:Reference number: Z20882; MUID:97129405; PMID:8973393

A:Accession: T30815

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1048 <NON>

A:Cross-references: EMBL:063926; NID:91572706; PID:91752707; PIDN:AAC60062.1

C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;

Query Match 40.0%; Score 6; DB 1; Length 1048;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGP 6
Db 652 SHLGP 657

RESULT 39

PFMSRB

platelet-derived growth factor receptor beta precursor - mouse

N:Contains: protein-tyrosine kinase (EC 2.7.1.112)

C:Species: Mus musculus (house mouse)

C:Date: 31-Mar-1993 #sequence-revision 31-Mar-1993 #text-change 10-Sep-1999

C:Accession: A25742

R:Yarden, Y.; Escobedo, J.A.; Kuang, W.J.; Yang-Feng, T.L.; Daniel, T.O.; Tremble, P.M.; Nature 323, 226-232, 1986

A:Title: Structure of the receptor for platelet-derived growth factor helps define a family

A:Reference number: A25742; MUID:87014762; PMID:3020426

A:Accession: A25742

A:Molecule type: mRNA

A:Residues: 1-1098 <YAR>

A:Cross-references: EMBL:X04367; NID:953618; PIDN:CAA27882.1; PID:953619

A:Note: part of this sequence, including the amino end of the mature protein, was confirmed by cDNA sequencing.

C:Comment: The extracellular domain is predicted to include five immunoglobulin-like domains.

C:Comment: The putified receptor was found to be ubiquitinated.

C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology; C:Keywords: ATP; autophosphorylation; glycoprotein; growth factor receptor; heterodimer; F:1-31/Domin: signal sequence #status predicted <SIG>

F:32-1098/Product: platelet-derived growth factor receptor beta #status predicted <MAT>

F:32-530/Domin: extracellular #status predicted <EXT>

F:46-101/Domin: immunoglobulin homology <IMN1>

F:141-191/Domin: immunoglobulin homology <IMN2>

F:227-292/Domin: immunoglobulin homology <IMN3>

F:428-509/Domin: immunoglobulin homology <IMN4>

F:531-554/Domin: transmembrane #status predicted <TMN>

F:555-1098/Domin: intracellular #status predicted <INT>

F:597-964/Domin: protein kinase homology <KIN>

F:605-613/Region: protein kinase ATP-binding motif

F:44,88,102,214,291,306,353,370,444,467,478/Binding site: carbohydrate (Asn) (covalent)

F:53-99,148-189,224-290,435-507/Disulfide bonds: #status predicted

F:633/Active site: Lys #status predicted

F:856/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 40.0%; Score 6; DB 1; Length 1098;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGP 6
Db 655 SHLGP 660

RESULT 40

PFHUCB

platelet-derived growth factor receptor beta precursor - human

N:Contains: protein-tyrosine kinase (EC 2.7.1.112)

C:Species: Homo sapiens (man)

C:Spec: Homo sapiens (man)

C:Date: 31-Dec-1992 #sequence-revision 31-Dec-1992 #text-change 11-Jun-1999

C:Accession: A28206; A31195; A38266; A31925; B31925; C31925

R:Gronwald, R.G.K.; Grant, F.J.; Haldeman, B.A.; Hart, C.E.; O'Hara, P.J.; Hagen, F.S.; Proc. Natl. Acad. Sci. U.S.A. 85, 3433-3439, 1988

```

A>Title: Cloning and expression of a cDNA coding for the human platelet-derived growth
A:Reference number: A28206; MWID:88217915; PMID:2835772
A:Accession: A28206
A:Molecule type: mRNA
A:Residues: 1-1106 <RGRO>
A:Cross-references: GB:J03278; NID:g189731; PIDN:AAA60049.1; PID:g189732
R:Classon-Welsh, L.; Eriksson, A.; Moren, A.; Severinsson, L.; Ek, B.; Oestman, A.;
Mol. Cell. Biol. 8, 3476-3486, 1988
A>Title: cDNA cloning and expression of a human platelet-derived growth factor (PDGF)
A:Reference number: A31195; MWID:89096941; PMID:2850496
A:Accession: A31195
A:Molecule type: mRNA
A:Residues: 1-240, 'D', 242-1106 <CIA>
A:Cross-references: GB:M21616; NID:g189729; PIDN:AAA36427.1; PID:g189730
R:Parthan, J.; Maekele, T.P.; Altalo, R.; Lehtvasialho, H.; Altalo, K.
Proc. Natl. Acad. Sci. U.S.A. 87, 8913-8917, 1990
A>Title: Putative tyrosine kinases expressed in K-562 human leukemia cells.
A:Reference number: A38268; MWID:91062389; PMID:2247464
A:Accession: A38268
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 828-884 <PAR>
R:Roberts, W.M.; Look, A.T.; Rousset, M.F.; Sherr, C.J.
Cell 55, 655-661, 1988
A>Title: Tandem linkage of human CSF-1 receptor (c-fms) and PDGF receptor genes.
A:Reference number: A90908; MWID:89028677; PMID:2846185
A:Accession: A31925
A>Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 676-727 <ROB>
A:Accession: B31925
A>Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 901-932 <RO2>
A:Accession: C31925
A>Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1047-1106 <RO3>
A:Comment: The extracellular domain is predicted to include five immunoglobulin-like
C:Genetics:
A:Gene: GDB:PDGFRB
A:Cross-references: GDB:120710; OMIM:173410
A:Map position: 5q31-5q32
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homolo
C:Keywords: ATP; autophosphorylation; glycoprotein; growth factor receptor; heterodim
F:1-33/Domain: signal sequence #status predicted <SIG>
F:33-1106/Product: platelet-derived growth factor receptor beta #status predicted <MA
F:33-531/Domain: extracellular #status predicted <EXT>
F:47-102/Domain: immunoglobulin homology <IM1>
F:142-193/Domain: immunoglobulin homology <IM2>
F:228-293/Domain: immunoglobulin homology <IM3>
F:429-510/Domain: immunoglobulin homology <IM4>
F:532-555/Domain: transmembrane #status predicted <TM>
F:556-1106/Domain: intracellular #status predicted <INT>
F:598-965/Domain: protein kinase homology <RIN>
F:606-614/Region: protein kinase ATP-binding motif
F:45-89,103,215,230,292,307,354,371,468,479/Binding site: carbohydrate (Asn) (covalen
F:54-100,149-190,235-261,436-508/disulfide bonds: #status predicted
F:634/Active site: Lys #status predicted
F:857/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predi

```

N:Alternate names: hypothetical protein T29F13.11
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 23-Mar-2001
 C:Accession: T01367; F84759
 R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul
 submitted to the EMBL Data Library, May 1998
 A:Description: Arabidopsis thaliana chromosome II BAC T29F13 genomic sequence.
 A:Reference number: 214179
 A:Accession: T01367
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1680 <R0U>
 A:Cross-references: EMBL:AC003096; NID:g3132469; PID:g3132477
 A:Experimental source: cultivar Columbia
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanden, S.E.; Umayan, L.; Tallon, L.
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487; PMID:10617197
 A:Accession: F84759
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1680 <SNO>
 A:Cross-references: GB:A002093; NID:g3132477; PID:NAC16266.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: T29F13.11; At2g34680
 A:Map position: 2
 A:Introns: 2713/3; 295/3; 317/3; 341/3; 363/3; 384/3; 406/3; 440/1; 474/1; 551/3; 600/2;
 ; 1372/2; 1433/1; 1485/1; 1529/1; 1548/3; 1578/3; 1622/3

Query Match 40.0%; Score 6; DB 2; Length 1680;
 Best Local Similarity 100.0%; Pred. No. 99;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 8 STPEER 13
 |||||
 Db 248 STPEER 253

RESULT 42
 A47392
 chromodomain-helicase-DNA-binding protein, CHD-1 - mouse
 N:Alternate names: K1BP protein
 C:Species: Mus musculus (house mouse)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: A47392; S21568
 R:Delmas, V.; Stokes, D.G.; Perry, R.P.
 Proc. Natl. Acad. Sci. U.S.A. 90, 2414-2418, 1993
 A:Title: A mammalian DNA-binding protein that contains a chromodomain and an SNF2/SWI2-1
 A:Reference number: A47392; MUID:93211972; PMID:8460153
 A:Accession: A47392
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1711 <DBL>
 A:Experimental source: S194 plasmacytoma cells
 A:Note: sequence inconsistent with the nucleotide translation
 A:Note: sequence extracted from NCBI backbone (NCBI:128272, NCBI:128273)
 R:Delmas, V.; Perry, R.P.
 submitted to the EMBL Data Library, May 1992
 A:Description: K1BP, a mammalian protein that contains the SNF2/SWI2 helicase domain als
 A:Reference number: S21568
 A:Accession: S21568
 A:Molecule type: mRNA
 A:Residues: 772-1711 <DE2>
 A:Cross-references: EMBL:X66028
 C:Superfamily: CHD-1 protein; chromobox homology
 C:Keywords: DNA binding
 F:293-336/Domain: chromobox homology <CB1>
 F:387-427/Domain: chromobox homology <CB2>

Query Match 40.0%; Score 6; DB 1; Length 1711;
 Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 6 HRSTPE 11
 |||||
 Db 1698 HRSTPE 1703

RESULT 43
 T28626
 variant-specific surface protein 2 - malaria parasite (Plasmodium falciparum)
 C:Species: Plasmodium falciparum
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
 C:Accession: T28626
 R:Su, X.Z.; Beattole, V.M.; Wertheimer, S.P.; Guinet, F.; Herfeldt, J.A.; Peterson,
 Cell 82, 89-100, 1995
 A:Title: The large diverse gene family var encodes proteins involved in cytoadherence
 A:Reference number: Z20487; MUID:95330813; PMID:7606788
 A:Accession: T28626
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-2664 <SUX>
 A:Cross-references: EMBL:L40609; NID:g886376; PID:g886378; PIDN:AA75398.1
 C:Genetics:
 A:Introns: 2197/3
 A:Note: var-2

Query Match 40.0%; Score 6; DB 2; Length 2664;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 7 RSTPE 12
 |||||
 Db 1335 RSTPE 1340

RESULT 44
 S58277
 insulin-like growth factor receptor type II - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 05-Nov-1999
 C:Accession: S58277
 R:Smrzka, O.W.; Stoger, R.; Kurzbauer, R.; Fag, I.; Fischer, G.F.; Barlow, D.P.
 submitted to the EMBL Data Library, January 1995
 A:Description: Conservation of a methylation imprint and a putative imprinting box at
 A:Reference number: S58277
 A:Accession: S58277
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-18 <SMR>
 A:Cross-references: EMBL:X83702; NID:g929644; PIDN:CA58675.1; PID:g929645
 C:Keywords: growth factor receptor

Query Match 33.3%; Score 5; DB 2; Length 18;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 11 ESRAA 15
 |||||
 Db 6 ESRAA 10

RESULT 45
 T04371
 thaumatin-like protein - barley (fragment)
 C:Species: Hordeum vulgare (barley)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 08-Oct-1999
 C:Accession: T04371
 R:Skadsen, R.W.; Herbst, J.M.
 submitted to the EMBL Data Library, July 1997
 A:Reference number: Z15316
 A:Accession: T04371
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA

A:Residues: 1-26 <SKA>
 A:Cross-references: EMBL:AF016328; NID:g2454603; PIDN:AAB71681.1; PID:g2454604
 A:Experimental source: cv. Morex
 C:Genetics:
 A:Gene: perm2

Query Match 33.3%; Score 5; DB 2; Length 26;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHR 7
 |||||
 DB 19 LGPHR 23

RESULT 46
 137554

homeobox - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 15-Oct-1999
 C:Accession: 137554
 R:Moretti, P.; Simmons, P.; Thomas, P.; Haylock, D.; Rathjen, P.; Vadas, M.; D'Andrea, F.
 Gene 144, 213-219, 1994
 A:Title: Identification of homeobox genes expressed in human haemopoietic progenitor cell
 A:Reference number: 137554; MUID:94314219; PMID:7518789
 A:Accession: 137554
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-39 <RES>
 A:Cross-references: EMBL:X74861; NID:g510992; PIDN:CAA52854.1; PID:g510993

C:Genetics:
 A:Gene: HPX-2
 C:Superfamily: unassigned homeobox proteins; homeobox homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation

Query Match 33.3%; Score 5; DB 2; Length 39;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESR 13
 |||||
 DB 32 TPESR 36

RESULT 47
 P86491

hypothetical protein CPJ0006 [imported] - Chlamydophila pneumoniae (strain J138)
 C:Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
 C:Accession: P86491
 R:Shital, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is
 Nucleic Acids Res. 28, 2311-2314, 2000
 A:Title: Comparison of whole genome sequences of Chlamydia pneumoniae J138.
 A:Reference number: A86491; MUID:20330349; PMID:10871362
 A:Accession: P86491
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-51 <STO>
 A:Cross-references: GB:BA000008; NID:g8978379; PIDN:BA98216.1; GSPDB:GN00142
 A:Experimental source: strain J138
 C:Genetics:
 A:Gene: CPJ0006

Query Match 33.3%; Score 5; DB 2; Length 51;
 Best Local Similarity 100.0%; Pred. No. 58;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
 |||||
 DB 29 RSTPE 33

RESULT 48

C72131
 hypothetical protein - Chlamydophila pneumoniae (strain CWL029)

C:Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
 C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 05-May-2000
 C:Accession: C72131
 R:Kaiman, S.; Mitchell, W.; Marathe, R.; Lammell, C.; Fan, J.; Olinger, L.; Grimwood,
 Nature Genet. 21, 385-389, 1999
 A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
 A:Reference number: A72000; MUID:99206606; PMID:10192388
 A:Accession: C72131
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-51 <ARN>
 A:Cross-references: GB:AE001585; GB:AE001363; NID:g4376255; PIDN:AAD18164.1; PID:g4376255
 A:Experimental source: strain CWL029
 C:Genetics:
 A:Gene: Cpn0006

Query Match 33.3%; Score 5; DB 2; Length 51;
 Best Local Similarity 100.0%; Pred. No. 58;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
 |||||
 DB 29 RSTPE 33

RESULT 49
 G72355

hypothetical protein - Thermotoga maritima (strain MSB8)
 C:Species: Thermotoga maritima
 C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
 C:Accession: G72355
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hic
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,
 C.M.
 Nature 399, 323-329, 1999
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome
 A:Reference number: A72200; MUID:99287316; PMID:10360571
 A:Accession: G72355
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-56 <ARN>
 A:Cross-references: GB:AE001735; GB:AE000512; NID:g4981122; PIDN:AAD35696.1; PID:g4981122
 A:Experimental source: strain MSB8
 C:Genetics:
 A:Gene: TM0611

Query Match 33.3%; Score 5; DB 2; Length 56;
 Best Local Similarity 100.0%; Pred. No. 63;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESR 13
 |||||
 DB 22 TPESR 26

RESULT 50
 Q0ECPT

ybeB protein - Escherichia coli (strain K-12)
 C:Species: Escherichia coli
 C>Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 01-Mar-2002
 C:Accession: A24995; C64798
 R:Asou, S.; Matsuzawa, H.; Ishino, F.; Strominger, J.L.; Matsubashi, M.; Ohta, T.
 Eur. J. Biochem. 160, 231-238, 1986
 A:Title: Nucleotide sequence of the ybeB gene and characteristics of the deduced amino
 A:Reference number: A91176; MUID:87030266; PMID:3533535
 A:Accession: A24995
 A:Molecule type: DNA
 A:Residues: 1-69 <ASO>
 A:Cross-references: GB:X04516; GB:D00001; GB:N00001; NID:g42313; PIDN:CAA28199.1; PID
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.,
 A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of *Escherichia coli* K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: C64798
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-69 <BLAT>
A:Cross-references: GB:AE000168; GB:U00096; NID:q1786849; PIDN:AMC73738.1; PID:q1786856;
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: ybeB
A:Map position: 15 min
C:Superfamily: *Escherichia coli* ybeB protein

Query Match 33.3%; Score 5; DB 1; Length 69;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRRA 15
Db 22 ESRRA 26

RESULT 51
C90713
hypothetical protein Ecs0675 [Imported] - *Escherichia coli* (strain O157:H7, substrain R1
C:Species: *Escherichia coli*
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 17-May-2002
C:Accession: C90713
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genc
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: C90713
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-69 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA834098.1; PID:q13360133; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain R1MD 0509952
C:Genetics:
A:Gene: Ecs0675
C:Superfamily: *Escherichia coli* ybeB protein

Query Match 33.3%; Score 5; DB 2; Length 69;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRRA 15
Db 22 ESRRA 26

RESULT 52
G85563
hypothetical protein ybeB [Imported] - *Escherichia coli* (strain O157:H7, substrain EDL93
C:Species: *Escherichia coli*
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 17-May-2002
C:Accession: G85563
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Hiller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: G85563
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-69 <STO>
A:Cross-references: GB:AE005174; NID:q12513538; PIDN:AA54971.1; GSPDB:GN00145; UWGP:Z07
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: ybeB
C:Superfamily: *Escherichia coli* ybeB protein

Query Match 33.3%; Score 5; DB 2; Length 69;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ESRRA 15
Db 22 ESRRA 26

RESULT 53
D72667
hypothetical protein APE5034 - *Aeropyrum pernix* (strain K1)
C:Species: *Aeropyrum pernix*
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C:Accession: D72667
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Ta
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, *Aero*
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: D72667
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-76 <KAW>
A:Cross-references: DDBJ:AP000060; NID:q5104188; PIDN:BAAT9740.1; PID:d1043526; PID:g
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE5034

Query Match 33.3%; Score 5; DB 2; Length 76;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPE 11
Db 11 RSTPE 15

RESULT 54
S28280
hypothetical protein B0464.3 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 09-Sep-1997
C:Accession: S28280
R:Kershaw, J.
submitted to the EMBL Data Library, December 1992
A:Reference number: S28278
A:Accession: S28280
A:Molecule type: DNA
A:Residues: 1-82 <KER>
A:Cross-references: EMBL:Z19152; NID:q6633; PID:q6636
C:Genetics:
A:Introns: 21/2; 51/3

Query Match 33.3%; Score 5; DB 2; Length 82;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHR 7
Db 62 LGPHR 66

RESULT 55
AD3436
hypothetical protein BME11474 [Imported] - *Brucella melitensis* (strain 16M)
C:Species: *Brucella melitensis*
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: AD3436
R:Delvecchio, V.G.; Kapral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanov
.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Let
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2003, 12:48:44 ; Search time 62 seconds

(without alignments)
32.238 Million cell updates/sec

Title: US-09-689-159a-2_COPY_346_360

Perfect score: 15

Sequence: 1 SHUGPMSRPESRAA 15

Scoring table: OLIGO

Searched: 908470 seqs, 133250620 residues

Word size: 0

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 200 summaries

Database: A.Geneseq_101002.*

```
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15	100.0	71	AAW05781	Presentinlin-1-1 res
2	15	100.0	164	AAW11787	Early onset Alzheimer
3	15	100.0	354	AAE12901	PS1 loop - Bacteri
4	15	100.0	407	AAW28507	Partial AD3 sequen
5	15	100.0	429	AAW41429	PS1/429 protein.
6	15	100.0	430	AAW05761	Presentinlin-1-2 D25
7	15	100.0	434	AAW05760	Presentinlin-1-1 D25
8	15	100.0	438	AAW05756	Presentinlin-1-1 del
9	15	100.0	463	AAW05734	Presentinlin-1-2. H.
10	15	100.0	463	AAW22948	Presentinlin-1 VRSQ

11	15	100.0	463	18	AAW12376	Human S182 gene pr
12	15	100.0	463	18	AAW11840	Early onset Alzheimer
13	15	100.0	463	19	AAW20854	Human presentinlin I
14	15	100.0	463	19	AAW23965	Human presentinlin-1
15	15	100.0	463	19	AAW42375	Human presentinlin I
16	15	100.0	465	21	AAW51393	Human S182 protein
17	15	100.0	465	21	AAW51394	Murine S182 protei
18	15	100.0	467	17	AAW05750	Presentinlin-1-1 C26
19	15	100.0	467	17	AAW05751	Presentinlin-1-1 P26
20	15	100.0	467	17	AAW05752	Presentinlin-1-1 P26
21	15	100.0	467	17	AAW05753	Presentinlin-1-1 E28
22	15	100.0	467	17	AAW05754	Presentinlin-1-1 A28
23	15	100.0	467	17	AAW05755	Presentinlin-1-1 L28
24	15	100.0	467	17	AAW05756	Presentinlin-1-1 G38
25	15	100.0	467	17	AAW05757	Presentinlin-1-1 L39
26	15	100.0	467	17	AAW05758	Presentinlin-1-1 L39
27	15	100.0	467	17	AAW05759	Presentinlin-1-1 C41
28	15	100.0	467	17	AAW05733	Presentinlin-1-1. H
29	15	100.0	467	17	AAW05735	Murine presentinlin.
30	15	100.0	467	17	AAW05736	Presentinlin-1-1 A79
31	15	100.0	467	17	AAW05737	Presentinlin-1-1 V82
32	15	100.0	467	17	AAW05738	Presentinlin-1-1 V06
33	15	100.0	467	17	AAW05739	Presentinlin-1-1 Y11
34	15	100.0	467	17	AAW05740	Presentinlin-1-1 M13
35	15	100.0	467	17	AAW05741	Presentinlin-1-1 M14
36	15	100.0	467	17	AAW05742	Presentinlin-1-1 M14
37	15	100.0	467	17	AAW05743	Presentinlin-1-1 H16
38	15	100.0	467	17	AAW05744	Presentinlin-1-1 L17
39	15	100.0	467	17	AAW05745	Presentinlin-1-1 G20
40	15	100.0	467	17	AAW05746	Presentinlin-1-1 I21
41	15	100.0	467	17	AAW05747	Presentinlin-1-1 I23
42	15	100.0	467	17	AAW05748	Presentinlin-1-1 A24
43	15	100.0	467	19	AAW05749	Presentinlin-1-1 A26
44	15	100.0	467	19	AAW23966	Human mutant S182
45	15	100.0	467	18	AAW27176	Human S182 gene, P
46	15	100.0	467	18	AAW11839	Human early onset
47	15	100.0	467	19	AAW56770	Human sapiens PS-1.
48	15	100.0	467	19	AAW23964	Human presentinlin-1.
49	15	100.0	467	19	AAW23965	Mouse presentinlin-1
50	15	100.0	467	19	AAW41430	PS1/467 protein.
51	15	100.0	467	19	AAW41431	Mouse PS1/467 prot
52	15	100.0	467	20	AAW24419	Human presentinlin-1
53	15	100.0	467	20	AAW23897	Mouse presentinlin-1
54	15	100.0	467	21	AAW07971	Amino acid sequenc
55	15	100.0	467	22	AAE10798	Amino acid sequenc
56	15	100.0	467	22	AAE63936	Human presentinlin-1
57	15	100.0	467	22	AAE05466	Human presentinlin (
58	15	100.0	467	22	AAE05563	Human presentinlin P
59	15	100.0	467	23	AAE05564	Human presentinlin P
60	15	100.0	467	23	AAO18049	Presentinlin protein
61	15	100.0	467	23	AAW79416	Human presentinlin-1
62	15	100.0	467	23	AAE17045	Human mutant prese
63	15	100.0	467	23	AAE17046	Human mutant prese
64	15	100.0	467	23	AAE17047	Human mutant prese
65	15	100.0	467	23	AAE17051	Human mutant prese
66	15	100.0	467	23	AAW05785	Presentinlin-1-1 res
67	15	100.0	467	23	AAW34093	Peptide derived fr
68	15	100.0	467	22	AAE12838	Presentinlin peptide
69	15	100.0	467	22	AAE12902	Presentinlin peptide
70	15	100.0	467	22	AAU58947	Propionibacterium
71	15	100.0	467	22	AAU58947	Zea mays protein f
72	15	100.0	467	22	AAU41658	Propionibacterium
73	15	100.0	467	22	AAU04044	Streptococcus coel
74	15	100.0	467	22	AAW97977	Human peptide #125
75	15	100.0	467	22	AAW97978	Human peptide #125
76	15	100.0	467	22	AAW38298	Peptide #5804 enco
77	15	100.0	467	22	AAW38298	Protein #5478 enco
78	15	100.0	467	22	AAW58920	Human brain expres
79	15	100.0	467	22	AAW71442	Human bone marrow
80	15	100.0	467	22	AAW19095	Peptide #3529 enco
81	15	100.0	467	22	AAW19095	Peptide #5771 enco
82	15	100.0	467	22	AAW19095	Human peptide enco
83	15	100.0	467	22	ABG41248	Novel human diagno

84	6	40.0	77	22	ABG06427	Novel human diagno
85	6	40.0	80	22	AAU42555	Proprionibacterium
86	6	40.0	80	22	ABG14485	Novel human diagno
87	6	40.0	84	22	ABG21779	Novel human diagno
88	6	40.0	85	22	ABG06213	Novel human diagno
89	6	40.0	89	22	ABG06199	Novel human diagno
90	6	40.0	89	22	ABG14499	Novel human diagno
91	6	40.0	89	22	ABG26397	Novel human diagno
92	6	40.0	93	22	ABG18825	Novel human diagno
93	6	40.0	96	20	AAU40019	Peptide sequence d
94	6	40.0	102	23	ABP35429	Human synthase-11k
95	6	40.0	103	22	AAU39273	Proprionibacterium
96	6	40.0	105	22	ABG11452	Novel human diagno
97	6	40.0	107	22	ABG11454	Novel human diagno
98	6	40.0	112	22	ABG06208	Novel human diagno
99	6	40.0	112	22	ABG06438	Novel human diagno
100	6	40.0	112	22	ABG14494	Novel human diagno
101	6	40.0	112	22	ABG28253	Novel human diagno
102	6	40.0	118	22	ABG24883	Novel human diagno
103	6	40.0	120	22	ABG27021	Novel human diagno
104	6	40.0	125	22	ABG14479	Novel human diagno
105	6	40.0	126	22	AAU22757	Novel human diagno
106	6	40.0	126	22	AAU94733	Human prostate can
107	6	40.0	134	22	AAU01111	Human reproductive
108	6	40.0	135	22	ABG06201	Human polypeptide
109	6	40.0	135	22	ABG14117	Novel human diagno
110	6	40.0	137	22	ABG14498	Novel human diagno
111	6	40.0	146	22	ABG11461	Novel human diagno
112	6	40.0	158	22	ABG28247	Novel human diagno
113	6	40.0	161	22	ABG06433	Novel human diagno
114	6	40.0	167	22	ABG19252	Novel human diagno
115	6	40.0	176	22	ABG14481	Novel human diagno
116	6	40.0	177	22	ABG06238	Novel human diagno
117	6	40.0	199	22	ABG28263	Novel human diagno
118	6	40.0	205	22	ABG28257	Novel human diagno
119	6	40.0	256	22	AAU65725	Novel human diagno
120	6	40.0	270	20	AAU43977	Lethal leaf spot p
121	6	40.0	303	22	ABG08486	Mouse protein kina
122	6	40.0	329	18	AAU24231	Novel human diagno
123	6	40.0	369	18	AAU31903	Human melanoma ass
124	6	40.0	411	18	AAU24229	Streptococcus pneu
125	6	40.0	415	22	AAU27706	Human melanoma ass
126	6	40.0	426	22	ABP27059	Human full-length
127	6	40.0	429	19	AAU70983	Streptococcus poly
128	6	40.0	429	21	AAU85136	Histidyl tRNA synt
129	6	40.0	429	22	AAU37913	Streptococcus pneu
130	6	40.0	475	18	AAU24228	Human melanoma ass
131	6	40.0	479	23	ABG97204	Novel human protei
132	6	40.0	513	21	AAU34332	zee mays protein f
133	6	40.0	525	18	AAU24230	Human melanoma ass
134	6	40.0	543	22	ABG04560	Novel human diagno
135	6	40.0	614	23	ABG93594	Hebldicidally activ
136	6	40.0	739	22	AAU93019	C glutamicum prote
137	6	40.0	795	23	ABP27401	Streptococcus poly
138	6	40.0	821	20	AAU28935	Platelet-derived g
139	6	40.0	853	22	AAU93437	Human polypeptide,
140	6	40.0	865	14	AAU30169	Marck's Disease Vi
141	6	40.0	1086	18	AAU36051	Hybrid Marek's dis
142	6	40.0	1089	22	ABG20753	Mouse ischaemic co
143	6	40.0	1098	23	ABG57338	Mouse ischaemic co
144	6	40.0	1106	10	AAU90646	Platelet-derived g
145	6	40.0	1106	10	AAU90127	Type B human plate
146	6	40.0	1106	13	AAU26205	Platelet-derived g
147	6	40.0	1106	17	AAU99690	Novel human diagno
148	6	40.0	1225	22	ABG03949	Novel human diagno
149	6	40.0	2034	20	AAU06300	Human activated ca
150	6	40.0	2129	22	ABG15478	Novel human diagno
151	6	40.0	2129	22	ABG20749	Novel human diagno
152	6	40.0	2139	22	ABG62393	Drosophila melanog
153	6	40.0	2353	20	AAU06299	Human activated ca
154	6	40.0	2353	20	ABG30840	Human voltage-depe
155	6	40.0	4545	18	AAU22611	Hybrid smc/tyl6 0

157	6	40.0	4550	18	AAU23716	Platenolide synth
158	6	40.0	4550	18	AAU22606	Platenolide synth
159	5	33.3	8	23	AAU47749	SIV Tat(28-35). S
160	5	33.3	8	23	AAU47751	CTL epitope T4t_SL
161	5	33.3	8	23	AAU47758	CTL epitope 94004
162	5	33.3	10	22	AAU88112	Saccharomyces cere
163	5	33.3	10	22	AAU88113	Saccharomyces cere
164	5	33.3	11	13	AAU29138	Mutation #10 of T7
165	5	33.3	11	17	AAU94612	Protein kinase bin
166	5	33.3	12	14	AAU34443	N-terminal sequenc
167	5	33.3	12	23	AAU01192	Human Brn-5 transc
168	5	33.3	13	18	AAU12798	Antigenic peptide
169	5	33.3	13	22	AAU52755	ApJ g protein-coup
170	5	33.3	14	19	AAU59108	EMV non-structura
171	5	33.3	14	19	AAU59109	EMV non-structura
172	5	33.3	15	15	AAU47671	HIV epitope #4. H
173	5	33.3	15	22	AAU78901	Tumour suppressor p
174	5	33.3	21	22	AAU01628	Human gene 28 enco
175	5	33.3	21	23	ABG63804	Human albumin fusi
176	5	33.3	34	22	ABG39746	Peptide #7252 enco
177	5	33.3	34	22	AAU60466	Human brain expres
178	5	33.3	34	22	AAU73110	Human bone marrow
179	5	33.3	34	22	AAU33328	Peptide #7365 enco
180	5	33.3	39	22	ABG42957	Human peptide enco
181	5	33.3	39	22	ABG22649	Novel human diagno
182	5	33.3	40	22	ABG27905	Novel human diagno
183	5	33.3	47	19	AAU79463	Staphylococcus aur
184	5	33.3	48	22	ABU40766	Peptide #8272 enco
185	5	33.3	48	22	AAU61626	Human brain expres
186	5	33.3	48	22	AAU74418	Human bone marrow
187	5	33.3	48	22	AAU34532	Peptide #8569 enco
188	5	33.3	50	21	AAU42306	Human OREX ORF2070
189	5	33.3	50	23	ABP05658	Human OREX protein
190	5	33.3	54	21	AAU57801	Arabidopsis thalia
191	5	33.3	54	23	ABP10197	Human OREX protein
192	5	33.3	55	22	AAU60387	Proprionibacterium
193	5	33.3	56	22	AAU99608	Human excretory re
194	5	33.3	58	22	AAU42423	Human kidney relat
195	5	33.3	58	22	AAU65001	Human 5' EST relat
196	5	33.3	58	22	ABG22141	Novel human diagno
197	5	33.3	59	22	AAU78786	Human protein SBO
198	5	33.3	59	22	AAU4591	Proprionibacterium
199	5	33.3	60	22	AAU40678	Proprionibacterium
200	5	33.3	60	22	AAU64335	Proprionibacterium

ALIGNMENTS

346-360
 pd6-28-95

RESULT 1	AAU05781	standard; peptide; 71 AA.
ID	AAU05781	
XX	AAU05781	
AC	AAU05781	
XX	AAU05781	
DF	28-JUL-1997	(first entry)
XX	28-JUL-1997	
DE	Presentin-1-1 residues 300-370.	
XX	Presentin-1-1 residues 300-370.	
KW	Presentin-1-1: human; hps1-2; PS-2; Integral membrane protein; AD; familial Alzheimer's disease; cerebral haemorrhage; schizophrenia; depression; antibody; gene expression modulator; therapy; muten.	
OS	Homo sapiens.	
PN	W0634099-A2.	
XX	W0634099-A2.	
PD	31-OCT-1996.	
XX	31-OCT-1996.	
PF	29-APR-1996;	96WO-CA00263.
XX	29-APR-1996;	
PR	31-JUL-1995;	95US-0509359.
PR	28-APR-1995;	95US-0431048.

XX 28-JUN-1995; 95US-0496841.
PA (HSCR-) HSC RES & DEV LP.
PA (UTOR) UNIV TORONTO GOVERNING COUNCIL.
PI Fraser PE, Rommens JM, St George-Hyslop PH;
XX WPT: 1996-497631/49.

DR
PT New presenilin genes - useful for diagnosis, therapy and drug
PT screening of familial Alzheimer's disease, cerebral disorders, etc.
XX
PS Claim 71; Page -, 178pp; English.

CC AAM05768-W05788 represent antigenic fragments of the human
CC presenilin-1 protein (see AAM05733 for wild type sequence). AAM05734
CC represents a different wild type form of presenilin-1 that results from
CC alternate splicing of the genomic DNA sequence. The presenilins are a
CC family of highly conserved integral membrane proteins with a common
CC structural motif, common alternate splicing patterns, and common
CC mutational hot spot regions. Mutations in PS genes are implicated in
CC familial Alzheimer's disease (AD) and possibly other diseases such as
CC cerebral haemorrhage, schizophrenia, depression etc., so detection of
CC mutations in the DNA encoding the wild type sequences can be used for
CC diagnosis of these diseases. The wild type proteins, or vectors that
CC express them or containing antisense sequences, antibodies selective for
CC these mutant forms of the proteins and modulators of PS gene expression
CC are potentially useful for treatment of AD etc. Transgenic animals are
CC useful as models for drug screening. The antibodies can also be used e.g.
CC for affinity purification and in immunoassays.

XX Sequence 71 AA:
SQ

Query Match 100.0%; Score 15; DB 17; Length 71;
Best Local Similarity 100.0%; Pred. No. 7 le-09;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

OY 1 SHLGPNRSTPESRAA 15
|||
DB 47 SHLGPNRSTPESRAA 61

RESULT 2
AAW11787
ID AAW11787 standard; Protein: 164 AA.

XX AAW11787;
AC
XX
DT 24-APR-1997 (first entry)
DE Early onset Alzheimer's disease gene product (activated T-cell).
XX
KM Early onset Alzheimer's disease; EOAD; diagnosis; therapy;
KW expressed sequence tag; EST; activated T cell.
XX
OS Homo sapiens.
XX

Key Location/Qualifiers
FT Misc-difference 144 /note= "amino acid residue at position 144 is
FT unidentified owing to degeneracy of the
FT nucleotide sequence"
FT Misc-difference 151 /note= "amino acid residue at position 151 is
FT unidentified owing to degeneracy of the
FT nucleotide sequence"
FT Misc-difference 160 /note= "amino acid residue at position 160 is
FT unidentified owing to degeneracy of the
FT nucleotide sequence"
PN W09701573-A2.

[illegible]

```

XX  Novel gamma secretase protein, useful in the production of amyloids, is
PT  capable of cleaving beta-amyloid precursor protein to produce beta
PT  amyloid peptide
XX
XX  Example 8; Page 125-126; 127pp; English.
PS
XX  The invention relates to the field of plaque amyloid deposits that are
CC  the hallmarks of Alzheimer's disease. In particular, the invention
CC  relates to an isolated, functionally-active protein that has
CC  gamma-secretase activity. Gamma-secretase activity is necessary for
CC  amyloid production. The present invention also relates to methods for
CC  isolating integral-membrane proteins and protein complexes, including
CC  the gamma-secretase protein of the invention. The method is useful for
CC  monitoring the cleavage of beta-amyloid precursor protein (betaAPP)
CC  by gamma-secretase. The present sequence is a fusion protein comprising
CC  bacterial glutathione-S-transferase (GST) and presenilin peptide, PSI
CC  loop. This sequence is used in the exemplification of the invention.
SQ  Sequence 354 AA;

Query Match          100.0%; Score 15; DB 22; Length 354;
Best Local Similarity 100.0%; Pred. No. 2.8e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY  1 SHLGPHRSTPESRAA 15
    |||
DB  291 SHLGPHRSTPESRAA 305

RESULT 4
AAM28507
XX  AAM28507 standard; Protein: 407 AA.
AC
XX  AAM28507;
XX
DT  07-DEC-1997 (first entry)
XX
DE  Partial AD3 sequence.
XX
AD3; AD4/AD3LP; Alzheimer's disease; chromosome; missegregation;
KW  presenilin; inhibitor; AD; trisomy 21.
XX
OS  Homo sapiens.
XX
FH  Key
FT  Location/Qualifiers
FT  Misc-difference 86
FT  /label= mutation
FT  /note= "M -> L"
FT  Misc-difference 103
FT  /label= mutation
FT  /note= "H -> R"
FT  Misc-difference 186
FT  /label= mutation
FT  /note= "A -> E"
FT  Misc-difference 226
FT  /label= mutation
FT  /note= "L -> V"
FT  Misc-difference 350
FT  /label= mutation
FT  /note= "C -> Y"
FT
XX  WO9707213-A2.
XX  27-FEB-1997.
XX
XX  15-AUG-1996; 96WO-US13314.
XX
XX  16-AUG-1995; 95US-0002448.
XX
XX  (HARD ) HARVARD COLLEGE.
XX
XX  Li J, Potter H;
PI

```

```

XX  WPI; 1997-165297/15.
DR  N-PSDB; AAT87402.
XX
XX  Identifying genes which cause chromosome missegregation - useful for
PT  identifying causes of and treatments for diseases, e.g. Alzheimer's
PT  disease, cancer and ageing
XX
XX  Disclosure; Fig 1; 77pp; English.
XX
XX  Identifying genes which cause improper chromosome segregation,
CC  screening for inhibitors of chromosome missegregation and processes
CC  caused by genes encoding chromosome missegregation promoters
CC  was exemplified using Alzheimer's disease. The sequences
CC  given in AAT87401 to AAT87426 can be used in the above methods.
CC  The five mutations indicated in the Features Table cosegregate
CC  with early-onset familial Alzheimer's disease. It is predicted
CC  that these mutations result in increased levels of cells with
CC  trisomy 21 in carriers of the mutation compared with non-carriers.
XX
SQ  Sequence 407 AA;

Query Match          100.0%; Score 15; DB 18; Length 407;
Best Local Similarity 100.0%; Pred. No. 3.2e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY  1 SHLGPHRSTPESRAA 15
    |||
DB  286 SHLGPHRSTPESRAA 300

RESULT 5
AAM41429
XX  AAM41429 standard; Protein: 429 AA.
AC
XX  AAM41429;
XX
DT  04-JUN-1998 (first entry)
XX
DE  PS1/429 protein.
XX
KW  Presenilin peptide; PS1/429; immunogen; immune response; PS1 gene;
KW  Alzheimer's disease; mitochondrial pathology; neurodegeneration;
XX  apoptosis.
XX
OS  Homo sapiens.
XX
FH  WO9746678-A1.
XX
XX  WO9746678-A1.
XX  11-DEC-1997.
XX
XX  03-JUN-1997; 97WO-US09272.
XX
XX  18-JUL-1996; 96US-0683315.
XX  06-JUN-1996; 96US-0659296.
XX
XX  (FARB ) BAYER CORP.
XX
XX  Chisholm JC, Davis JN, Drache B;
XX  WPI; 1998-042186/04.
XX  N-PSDB; AAV17357.
XX
XX  DNA encoding presenilin peptide PS1/429 and its analogues - useful
PT  for diagnosis and treatment of Alzheimer's disease
XX
XX  Claim 24; Fig 1; 77pp; English.
XX
XX  This sequence is the PS1/429 presenilin peptide (II) of the
CC  invention. Cells transformed with the DNA are used to produce recombinant
CC  (II) and analogues, useful e.g. as immunogens for generating an immune
CC  response against PS1/429. (II) is a new product of the PS1 gene,
CC  mutations in which cause Alzheimer's disease (AD). The nucleic acids are

```


CC highly conserved integral membrane proteins with a common structural motif, common alternate splicing patterns, and common mutational hot spot regions. Mutations in PS genes are implicated in familial Alzheimer's disease (AD) and possibly other diseases such as cerebral haemorrhage, schizophrenia, depression etc., so detection of mutations in the DNA encoding the wild type sequences can be used for diagnosis of these diseases. The wild type proteins, or vectors that express them or containing antisense sequences, antibodies selective for these mutant CC forms of the proteins and modulators of PS gene expression are CC potentially useful for treatment of AD etc. Transgenic animals are CC useful as models for drug screening. The antibodies can also be used e.g. CC for affinity purification and in immunoassays.

SQ Sequence 434-AA;

Query Match 100.0%; Score 15; DB 17; Length 434;
Best Local Similarity 100.0%; Pred. No. 3.4e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLPHRSTPESRAA 15
|||||

Db 313 SHLPHRSTPESRAA 327

RESULT 8
AAW05756
ID AAW05756 standard; Protein: 438 AA.
XX AAW05756;

AC AAW05756;

DT 23-JUL-1997 (first entry)

DE Presentin-1-1 delta291-319 mutation.

XX Presentin-1; human; hps1-1; hps1-2; PS-2; Integral membrane protein; AD;
KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
KM depression; antibody; gene expression modulator; therapy; mutein.
XX Homo sapiens.
XX

OS Homo sapiens.

XX

FT Key Location/Qualifiers
FT Misc-difference 290..291
FT /note="site of 29 residue deletion"

PN W09634099-A2.
XX
XX 31-OCT-1996.
XX
XX 29-APR-1996; 96WO-CA00263.
XX
XX 31-JUL-1995; 95US-0509359.
PR 28-APR-1995; 95US-0431048.
PR 28-JUN-1995; 95US-0496841.
XX
XX (HSCR-) HSC RES & DEV LP.
PA (UTOR) UNIV TORONTO GOVERNING COUNCIL.
XX
XX Fraser PE, Rommens JM, St George-Hyslop PH;
XX WPI; 1996-497631/49.
XX
XX New presentin genes - useful for diagnosis, therapy and drug
XX screening of familial Alzheimer's disease, cerebral disorders, etc.
XX
XX Claim 3; Page -; 178pp; English.

PS AAW05736-W05760 represent mutated versions of the human presentin-1-1
CC protein (see AAW05733 for wild type sequence). AAW05734 represents a
CC different wild type form of presentin-1 that results from alternate
CC splicing of the genomic DNA sequence. The presentins are a family of
CC highly conserved integral membrane proteins with a common structural
CC motif, common alternate splicing patterns, and common mutational hot
CC spot regions. Mutations in PS genes are implicated in familial

CC Alzheimer's disease (AD) and possibly other diseases such as cerebral haemorrhage, schizophrenia, depression etc., so detection of mutations in the DNA encoding the wild type sequences can be used for diagnosis of these diseases. The wild type proteins, or vectors that express them or containing antisense sequences, antibodies selective for these mutant CC forms of the proteins and modulators of PS gene expression are CC potentially useful for treatment of AD etc. Transgenic animals are CC useful as models for drug screening. The antibodies can also be used e.g. CC for affinity purification and in immunoassays.

SQ Sequence 438 AA;

Query Match 100.0%; Score 15; DB 17; Length 438;
Best Local Similarity 100.0%; Pred. No. 3.4e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLPHRSTPESRAA 15
|||||

Db 317 SHLPHRSTPESRAA 331

RESULT 9
AAW05734
ID AAW05734 standard; Protein: 463 AA.
XX AAW05734;
XX AAW05734;

AC AAW05734;

DT 23-JUL-1997 (first entry)

DE Presentin-1-2.

XX Presentin-1; human; hps1-1; hps1-2; PS-2; Integral membrane protein; AD;
KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
KM depression; antibody; gene expression modulator; therapy.
XX Homo sapiens.
XX
XX W09634099-A2.
XX
XX 31-OCT-1996.
XX
XX 29-APR-1996; 96WO-CA00263.
XX
XX 31-JUL-1995; 95US-0509359.
PR 28-APR-1995; 95US-0431048.
PR 28-JUN-1995; 95US-0496841.
XX
XX (HSCR-) HSC RES & DEV LP.
PA (UTOR) UNIV TORONTO GOVERNING COUNCIL.
XX
XX Fraser PE, Rommens JM, St George-Hyslop PH;
XX WPI; 1996-497631/49.
XX N-PSDB; AAT40029.
XX
XX New presentin genes - useful for diagnosis, therapy and drug
XX screening of familial Alzheimer's disease, cerebral disorders, etc.
XX
XX Claim 2; Page 132-134; 178pp; English.

PS AAW05733 and AAW05734 represent the two different forms of wild type
CC human presentin-1 (PS-1). This form of presentin-1 results from
CC alternate splicing of the genomic DNA sequence. AAW05762 represents the
CC coding sequence for wild type human PS-2. The presentins are a family
CC of highly conserved integral membrane proteins with a common structural
CC motif, common alternate splicing patterns, and common mutational hot
CC spot regions. Mutations in PS genes are implicated in familial
CC Alzheimer's disease (AD) and possibly other diseases such as cerebral
CC haemorrhage, schizophrenia, depression etc., so detection of mutations in
CC the DNA encoding these sequences can be used for diagnosis of these
CC diseases. These proteins, or vectors that express them or containing
CC antisense sequences, antibodies selective for mutant forms of these
CC proteins (such as AAW05736) and modulators of PS gene expression are

CC potentially useful for treatment of AD etc. Transgenic animals are useful
CC as models for drug screening. The antibodies can also be used e.g. for
CC affinity purification and in immunoassays.

XX Sequence 463 AA;

Query Match 100.0%; Score 15; DB 17; Length 463;
Best Local Similarity 100.0%; Pred. No. 3.6e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPBRSPEPESRAA 15
DB 342 SHLGPBRSPEPESRAA 356

RESULT 10
AAW22948
ID AAW22948 standard; Protein; 463 AA.

XX AAW22948;

DT 19-MAR-1998 (first entry)

XX Presentin-1 VR5Q variant.

XX Identification: determination; neurological disease susceptibility;
KW detection; alternative splice site; polyadenylated mRNA transcript;
KM familial Alzheimer's disease; FAD; presentin 1; VR5Q variant.

OS Homo sapiens.

PN EP791660-A1.

XX 27-APR-1997.

XX 14-FEB-1997; 97EP-0300988.

PR 22-FEB-1996; 96US-0012077.

PA (SMK-) SMITHKLINE BEECHAM CORP.
PA (UYSE-) UNIV SOUTH FLORIDA DEPT PSYCHIATRY.
PA (UNIM) UNIV WASHINGTON.

PI Barton A, Goate A, Hardy J;

DR WPI: 1997-418049/39.

DR N-PSDB; AAT75576.

PT Diagnosis of, or susceptibility to neurological disease -
PT specifically Alzheimer's disease, by detecting aberrant splicing in
PT mRNA

PS Example 1; Pages 9-11; 21pp; English.

XX Identifying susceptibility to a neurological disease, comprises
CC detecting an alternative splice site in a polyadenylated mRNA
CC transcript in a sample of genetic material, where the alternative
CC splice site encodes AAW22944, or detecting AAW22944 in the protein
CC encoded by the mRNA. Tests on 3 early onset sporadic Alzheimer's
CC disease (FAD) patients, 6 late onset sporadic Alzheimer's disease
CC (AD) patients and 4 neurologically normal subjects, indicated that
CC mRNA transcripts of the presentin 1 gene in samples from various
CC brain regions occur in 2 forms, PS-1-long (containing a VR5Q
CC motif) and PS-1-short (lacking the VR5Q motif, i.e. the protein
CC denoted by the present sequence), and that the PS-1-long levels
CC in hippocampus and frontal cortex samples are significantly lower
CC in FAD patients than in AD and normal subjects.

XX Sequence 463 AA;

Query Match 100.0%; Score 15; DB 18; Length 463;
Best Local Similarity 100.0%; Pred. No. 3.6e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPBRSPEPESRAA 15
DB 342 SHLGPBRSPEPESRAA 356

RESULT 11
AAW12376
ID AAW12376 standard; Protein; 463 AA.

XX AAW12376;

DT 17-JUN-1997 (first entry)

XX Human S182 gene product.

XX S182 gene; familial Alzheimer's disease; diagnosis;
XX transgenic animal.

OS Homo sapiens.

PN WO9703999-A1.

XX 06-FEB-1997.

XX 26-JUN-1996; 96MO-US11065.

XX 02-APR-1995; 95US-0001800.

XX 18-JUL-1995; 95US-0001800.

PA (UYSE-) UNIV SOUTH FLORIDA.
PA (UNIM) UNIV WASHINGTON SCHOOL MED.

PI Goate AM, Hardy JA;

DR WPI: 1997-132571/12.

DR N-PSDB; AAT63207.

PT New mutants of the S182 gene associated with familial Alzheimer's
PT disease - and related protein and transgenic animals, useful as
PT models for screening and assessing potential drugs

PS Disclosure; Fig 1A-D; 26pp; English.

XX A polypeptide (AAW12376) is the product of an S182 gene cDNA clone
CC (AAT63207) isolated from a human brain library. Several mutations in
CC the S182 gene have been found in families with members affected by
CC early onset Alzheimer's disease (AD): in 2 families Met to Val at
CC position 135; in 3 families Met to Val at 142; in 1 family Pro to
CC Ser at 263; in 4 families Glu to Ala at 276; and in 1 family Glu to
CC Gly at 280. Detection of the mutations is used to diagnose AD, or
CC a predisposition to it. Transgenic animals can be produced that
CC are useful as models for screening and assessing potential drugs.

XX Sequence 463 AA;

Query Match 100.0%; Score 15; DB 18; Length 463;
Best Local Similarity 100.0%; Pred. No. 3.6e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPBRSPEPESRAA 15
DB 342 SHLGPBRSPEPESRAA 356

RESULT 12
AAW11840
ID AAW11840 standard; Protein; 463 AA.

XX AAW11840;

DT 07-MAY-1997 (first entry)

XX

DE Early onset Alzheimer's disease (EOAD) splice variant polypeptide.
 XX
 KW Early onset Alzheimer's disease; EOAD; neurodegenerative disease;
 KW diagnosis; therapy; inhibitor; antagonist; antibody.
 XX
 OS Homo sapiens.
 XX MO9703086-A1.
 PN 30-JUN-1997.
 XX
 XX 26-JUN-1996; 96WO-US11064.
 PF
 XX 18-JUL-1995; 95US-0001501.
 PR 13-JUL-1995; 95US-0001142.
 XX
 PA (UYSE-) UNIV SOUTH FLORIDA.
 XX
 PI Hardy JA;
 XX
 DR WPI: 1997-118980/11.
 DR N-PSDB; AAT59536.
 XX
 PT Early onset Alzheimer's disease gene - useful for diagnosing a
 PT pre-disposition to Alzheimer's disease
 XX
 PS Disclosure: Fig 2; 44pp; English.
 XX
 CC A 463-amino acid polypeptide (AAW11840) is the product of a full-
 CC length cDNA (AAT59536) of an early onset Alzheimer's disease (EOAD)
 CC splice variant gene. A 467-amino acid polypeptide (AAW11839) is the
 CC product of a full-length cDNA (AAT59535) of the EOAD gene. The 2
 CC polypeptides can be produced in transformed host cells and used to
 CC raise antibodies, or to identify antagonist/inhibitor cpds. useful
 CC in the treatment of Alzheimer's disease, esp. EOAD.
 CC
 XX
 SQ Sequence 463 AA:
 Query Match 100.0%; Score 15; DB 18; Length 463;
 Best Local Similarity 100.0%; Pred. No. 3.6e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SHLGPHRSTPESRAA 15
 DB 342 SHLGPHRSTPESRAA 356
 RESULT 13
 AAW20854
 ID AAW20854 standard; Protein; 463 AA.
 AC AAW20854;
 XX
 DT 22-JUL-1999 (first entry)
 XX
 DE Human presenilin I wild type protein fragment.
 XX
 KW Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
 KW frameshift mutation; age-related disease; neurodegenerative disorder;
 KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGp-C; NGP-A;
 KW high mobility group protein-C; neuroendocrine specific protein A.
 XX
 OS Homo sapiens.
 XX
 PN MO9845322-A2.
 XX
 PD 15-OCT-1998.

XX
 PF 02-APR-1998; 98WO-IB00705.
 XX
 PR 10-APR-1997; 97US-0043163.
 XX
 PA (UYUT-) RIJKSUNIV UTRECHT.
 PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
 PA (UYRO-) UNIV ROTTERDAM ERASMUS.
 XX
 PI Burbach JPH, Grosveld FG, Van Leeuwen FW;
 XX
 DR WPI: 1998-609901/51.
 DR N-PSDB; AAX75761.
 XX
 PT Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 PT RNA
 XX
 PS Disclosure: Figure 10; 258pp; English.
 XX
 CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-APP), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGp-C) and neuroendocrine specific protein A.
 CC
 XX
 SQ Sequence 463 AA:
 Query Match 100.0%; Score 15; DB 19; Length 463;
 Best Local Similarity 100.0%; Pred. No. 3.6e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SHLGPHRSTPESRAA 15
 DB 342 SHLGPHRSTPESRAA 356
 RESULT 14
 AAW23965
 ID AAW23965 standard; Protein; 463 AA.
 AC AAW23965;
 XX
 DT 20-JUL-1998 (first entry)
 XX
 DE Human presenilin-1.
 XX
 KW Presenilin-1; psi gene; human; familial Alzheimer's disease; FAD;
 KW cerebral haemorrhage; schizophrenia; depression; epilepsy;
 KW mental retardation; diagnosis; therapy; transgenic animal.
 XX
 OS Homo sapiens.
 XX
 PN Key Location/Qualifiers
 FT Misc-difference 177
 FT /note= "Phe17Ser mutation site (Claim 1)"
 FT Misc-difference 439
 FT /note= "Ile439Val mutation site (Claim 1)"


```

FT      Misc-difference 253
FT      /note= "Asp2537Ala mutation site, associated with
FT      FT      Misc-difference 254..286
FT      FT      /note= "residue 254-286 deletion mutant, associated
FT      FT      with Asp253Ala mutation (Claim 1)"
XX
XX      WO801549-A2.
XX      15-JAN-1998.
XX      04-JUL-1997; 97WO-CA00475.
XX      02-JAN-1997; 97US-0034590.
XX      PR      05-JUL-1996; 96US-0021673.
XX      PR      12-JUL-1996; 96US-0021700.
XX      PR      08-NOV-1996; 96US-0029895.
XX      (HSCR-) HSC RES & DEV LP.
XX      (UTOR ) UNITV TORONTO GOVERNING COUNCIL.
XX      Fraser PE, Rommens JM, St George-Hyslop PH;
XX      PI      WPI: 1998-286355/25.
XX      DR      N-PSDB: AAV04667.
XX      New isolated mutant presenilin-1 genes - useful for developing
XX      PT      products for use in detection, diagnosis and therapy of Alzheimer's
XX      PT      disease and for drug screening
XX      Claim 1; Page 165-186; 238pp; English.
XX
XX      This polypeptide comprises human presenilin-1 (hps1). Its amino
XX      CC      acid sequence was deduced from an isolated cDNA clone (see AAV04667).
XX      CC      Another hps1 sequence (see AAM23964) results from alternative
XX      CC      splicing of the hps1 mRNA transcript. A murine ps1 homologue (see
XX      CC      AAM23966) and a human presenilin-2 protein (see AAM23967) are also
XX      CC      provided. Mutations in the ps-1 and ps-2 genes are linked to the
XX      CC      development in humans of forms of familial Alzheimer's disease
XX      CC      (FAD) and may be causative of other disorders, e.g. cognitive,
XX      CC      intellectual, neurological or physiological disorders such as
XX      CC      cerebral haemorrhage, schizophrenia, depression, mental retardation
XX      CC      and epilepsy. Use of the nucleic acids and proteins comprising or
XX      CC      derived from the presenilins is made in screening and diagnosing
XX      CC      FAD, identifying and developing therapeutics for treatment of FAD,
XX      CC      and in producing cell lines and transgenic animals useful as models
XX      CC      of FAD. Methods for identifying substances that bind to, or
XX      CC      modulate the activity of a presenilin protein, and methods for
XX      CC      identifying substances that affect the interaction of a
XX      CC      presenilin-interacting protein with a presenilin protein are also
XX      CC      disclosed.
XX      Sequence 463 AA;
XX
XX      Query Match 100.0%; Score 15; DB 19; Length 463;
XX      Best Local Similarity 100.0%; Pred. NO. 3.0e-08;
XX      Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX      1 SHLGPQRSTPESRAA 15
XX      |||
XX      342 SHLGPQRSTPESRAA 356
XX
XX      RESULT 15
XX      AAM42375
XX      AAM42375 standard; Protein: 463 AA.
XX
XX      AAM42375:
XX
XX      08-JUN-1998 (first entry)
XX
XX      Human presenilin 1 gene product.
XX

```

KW	Presentin I gene; familial adult onset Alzheimer's disease; FAD;
XX	diagnosis; marker; VRXQ motif; human.
XX	
OS	Homo sapiens.
XX	
PN	EPH4.157-A2.
XX	
PD	29-DEC-1997.
XX	
PF	17-JUN-1997; 97EP-0304249.
XX	
PR	18-JUN-1996; 96US-0019991.
XX	
PA	(SMIK) SMITHKLINE BEECHAM CORP.
XX	(DYSE-) UNIV SOUTH FLORIDA.
PA	(UNIM) UNIV WASHINGTON.
XX	
PI	Barton ADL, Goate AM, Hardy J;
XX	
DR	WPI: 1998-044335/05.
XX	N-PSDB; AA003246.
XX	
PT	DNA encoding presenelin I - useful for diagnosis of Alzheimer's
PT	disease, drug screening, etc.
XX	
PS	Claim 4; Page 13-14; 20pp; English.
XX	
CC	This protein comprises the 463-amino acid human presenelin I
CC	(PS-I) gene product. A novel variant of PS-1 encoded by a clone
CC	isolated from a human cerebellar cDNA library contains a 4-amino
CC	acid insertion (VRSQ) between residues 26 and 27. This variant
CC	arises from alternative use of a 5' exon donor site in the exon
CC	3/intron 3 boundary of the PS-1 gene (see also AA003246). The
CC	4-amino acid motif can be used as a diagnostic marker for variants
CC	of presenelin genes associated with Alzheimer's disease and familial
CC	adult onset Alzheimer's disease (FAD). Methods are provided for
CC	detecting the presence or absence of a 4-amino acid motif (VRXQ,
CC	where X is a hydrophilic amino acid) in expressed proteins that
CC	arise from aberrant alternative splicing of pre-mRNA in genes
CC	associated with normal neurological function, which are useful for
CC	detecting neurodegenerative disease. The presence of these
CC	variants suggest that mutational events have occurred. Methods to
CC	measure the levels of gene expression of such genes to detect
CC	neurodegenerative diseases are provided. Nucleotide sequences and
CC	intron-exon junctional sequences of examples of this splicing
CC	variant and probes (see AA003247-49) for detecting this variant which
CC	are useful as diagnostic reagents are also provided.
XX	
XX	
SQ	Sequence 463 AA:
	Query Match 100.0%; Score 15; DB 19; Length 463;
	Best Local Similarity 100.0%; Pred. No. 3.6e-08;
	Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 SHLGPHRSTPESRAA 15
Db	342 SHLGPHRSTPESRAA 356
RESULT 16	
AAVS1393	
ID	AAVS1393 standard; Peptide; 465 AA.
XX	
AC	AAVS1393;
XX	
DT	04-MAY-2000 (first entry)
XX	
DE	Human S182 protein.
XX	
KW	SPE-4; immunogen; diagnosis; nematode; Alzheimer's disease; brain;
KW	muscle; peripheral blood cell; neuroprotectant; nootropic; S182; human.
OS	Homo sapiens.

[illegible]

XX	Example 4; Column 35-40; 23pp; English.
PS	
XX	This invention describes a novel immunogenic composition comprising at
CC	least 1 antigenic component selected from an antigenic peptide (I)
CC	(linked to a carrier) or a multiantigenic peptide (II). The composition
CC	is useful for diagnosing or monitoring SPE-4 related protein profiles of
CC	nematodes and/or Alzheimer's disease patients, either in postmortem
CC	tissue, or from other tissue samples, where the tissue is from the brain,
CC	muscle or peripheral blood cells. The immunogenic composition can be
CC	used to diagnose Alzheimer's noninvasively and has neuroprotective and
CC	neurotropic activity. This sequence represents the murine S182 protein
CC	which is used in the method of the invention.
XX	
SQ	Sequence 465 AA:
XX	
Query Match	100.0%; Score 15; DB 21; Length 465;
Best Local Similarity	100.0%; Pred. No. 3.6e+08;
Matches 15; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	1 SHLGPHRSPESRAA 15
Dd	344 SHLGPHRSTPESRAA 358
RESULT 18	
AAM05750	
ID	AAM05750 standard; Protein; 467 AA.
XX	
AC	AAM05750;
XX	
DT	23-JUL-1997 (first entry)
XX	
DE	Presentin-1-1 C263K mutation.
XX	
KW	Presentin-1; human; hPS1-1; hPS1-2; PS-2; integral membrane protein; AD;
KM	familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
KW	depression; antibody; gene expression modulator; therapy; muteln.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Modified-site Location/Qualifiers
FT	263
FT	/label= C263K
XX	
PN	WO934099-A2.
XX	
PD	31-OCT-1996.
XX	
PE	29-APR-1996; 96WO-CA00263.
XX	
PR	31-JUL-1995; 95US-0509359.
XX	
PR	28-APR-1995; 95US-0431048.
XX	
PR	28-JUN-1995; 95US-0496841.
XX	
PA	(HSCR-) HSC RES & DEV LP.
PA	(UTOR) UNIV TORONTO GOVERNING COUNCIL.
PI	
PL	Fraser PE, Rommens JM, St George-Hyslop PH;
XX	
DR	WPI; 1996-497631/49.
XX	
PT	New presentin genes - useful for diagnosis, therapy and drug
XX	screening of familial Alzheimer's disease, cerebral disorders, etc.
XX	
PS	Claim 3; Page -; 178pp; English.
XX	
AA	AAM05736-W05760 represent mutated versions of the human presentin-1-1
CC	protein (see AAM05733 for wild type sequence). AAM05734 represents a
CC	different wild type form of presentin-1 that results from alternate
CC	splicing of the genomic DNA sequence. The presentins are a family of
CC	highly conserved integral membrane proteins with a common structural
CC	motif, common alternate splicing patterns, and common mutational hot

CC spot regions. Mutations in PS genes are implicated in familial
CC Alzheimer's disease (AD) and possibly other diseases such as cerebral
CC haemorrhage, schizophrenia, depression etc., so detection of mutations in
CC the DNA encoding the wild type sequences can be used for diagnosis of
CC these diseases. The wild type proteins, or vectors that express them or
CC containing antisense sequences, antibodies selective for these mutant
CC forms of the proteins and modulators of PS gene expression are
CC potentially useful for treatment of AD etc. Transgenic animals are
CC useful as models for drug screening. The antibodies can also be used e.g.
CC for affinity purification and in immunoassays.

XX Sequence 467 AA;

Query Match 100.0%; Score 15; DB 17; Length 467;

Best Local Similarity 100.0%; Pred. No. 3,6e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGPSTPESRAA 15
Db 346 SHLGPSTPESRAA 360

RESULT 19

AAW05751 ID AAW05751 standard; Protein: 467 AA.

XX AAW05751:

XX 23-JUL-1997 (first entry)

XX Presenilin-1-1 P264L mutation.

XX Presenilin-1; human: hPS1-1; hPS1-2; PS-2; Integral membrane protein; AD;
XX familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
XX depression; antibody; gene expression modulator; therapy; mutain.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Modified-site 264 /Label= P264L

XX W09634099-A2.

XX 31-OCT-1996.

XX 29-APR-1996; 96WO-CA00263.

XX 31-JUL-1995; 95US-0509359.

XX 28-APR-1995; 95US-0431048.

XX 28-JUN-1995; 95US-0496841.

XX (HSCR-) HSC RES & DEV LP.
XX (UTOR) UNIV TORONTO GOVERNING COUNCIL.

XX Fraser PE, Rommens JM, St George-Hyslop PH;
XX WPI; 1996-497631/49.

XX New presenilin genes - useful for diagnosis, therapy and drug
XX screening of familial Alzheimer's disease, cerebral disorders, etc.

XX Claim 3; Page -: 178pp; English.

XX AAW05736-W05760 represent mutated versions of the human presenilin-1-1
XX protein (see AAW05733 for wild type sequence). AAW05734 represents a
XX different wild type form of presenilin-1 that results from alternate
XX splicing of the genomic DNA sequence. The presenilins are a family of
XX highly conserved integral membrane proteins with a common structural
XX motif, common alternate splicing patterns, and common mutational hot
XX spot regions. Mutations in PS genes are implicated in familial
XX Alzheimer's disease (AD) and possibly other diseases such as cerebral
XX haemorrhage, schizophrenia, depression etc., so detection of mutations in

CC the DNA encoding the wild type sequences can be used for diagnosis of
CC these diseases. The wild type proteins, or vectors that express them or
CC containing antisense sequences, antibodies selective for these mutant
CC forms of the proteins and modulators of PS gene expression are
CC potentially useful for treatment of AD etc. Transgenic animals are
CC useful as models for drug screening. The antibodies can also be used e.g.
CC for affinity purification and in immunoassays.

XX Sequence 467 AA;

Query Match 100.0%; Score 15; DB 17; Length 467;

Best Local Similarity 100.0%; Pred. No. 3,6e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGPSTPESRAA 15
Db 346 SHLGPSTPESRAA 360

RESULT 20

AAW05752 ID AAW05752 standard; Protein: 467 AA.

XX AAW05752:

XX 23-JUL-1997 (first entry)

XX Presenilin-1-1 P267S mutation.

XX Presenilin-1; human: hPS1-1; hPS1-2; PS-2; Integral membrane protein; AD;
XX familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
XX depression; antibody; gene expression modulator; therapy; mutain.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Modified-site 267 /Label= P267S

XX W09634099-A2.

XX 31-OCT-1996.

XX 29-APR-1996; 96WO-CA00263.

XX 31-JUL-1995; 95US-0509359.

XX 28-APR-1995; 95US-0431048.

XX 28-JUN-1995; 95US-0496841.

XX (HSCR-) HSC RES & DEV LP.
XX (UTOR) UNIV TORONTO GOVERNING COUNCIL.

XX Fraser PE, Rommens JM, St George-Hyslop PH;
XX WPI; 1996-497631/49.

XX New presenilin genes - useful for diagnosis, therapy and drug
XX screening of familial Alzheimer's disease, cerebral disorders, etc.

XX Claim 3; Page -: 178pp; English.

XX AAW05736-W05760 represent mutated versions of the human presenilin-1-1
XX protein (see AAW05733 for wild type sequence). AAW05734 represents a
XX different wild type form of presenilin-1 that results from alternate
XX splicing of the genomic DNA sequence. The presenilins are a family of
XX highly conserved integral membrane proteins with a common structural
XX motif, common alternate splicing patterns, and common mutational hot
XX spot regions. Mutations in PS genes are implicated in familial
XX Alzheimer's disease (AD) and possibly other diseases such as cerebral
XX haemorrhage, schizophrenia, depression etc., so detection of mutations in
XX the DNA encoding the wild type sequences can be used for diagnosis of
XX these diseases. The wild type proteins, or vectors that express them or
XX containing antisense sequences, antibodies selective for these mutant

CC forms of the proteins and modulators of PS gene expression are
 CC potentially useful for treatment of AD etc. Transgenic animals are
 CC useful as models for drug screening. The antibodies can also be used e.g.
 CC for affinity purification and in immunoassays.

XX Sequence 467 AA;

Query Match 100.0%; Score 15; DB 17; Length 467;
 Best Local Similarity 100.0%; Pred. No. 3.6e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPHRSTPESRAA 15
 |||||
 DB 346 SHLGPHRSTPESRAA 360

RESULT 21

AAW05753 ID AAW05753 standard; Protein; 467 AA.

XX AAW05753;

DT 23-JUL-1997 (first entry)

DE Presenilin-1-1 E280A/G mutation.

KW Presenilin-1; human; hps1-1; hps1-2; PS-2; Integral membrane protein; AD;
 KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
 KW depression; antibody; gene expression modulator; therapy; mutin.

OS Homo sapiens.

FH Key Location/Qualifiers
 FT Modified-site 280
 FT /label= E280X
 FT /note= "X = Ala, Gly"

PN WO9634099-A2.

XX 31-OCT-1996.

PF 29-APR-1996; 96WO-CA00263.

XX 31-JUL-1995; 95US-0509359.

PR 28-APR-1995; 95US-0431048.

XX 28-JUN-1995; 95US-0496841.

PA (HSCR-) HSC RES & DEV LP.
 PA (UTOR) UNIV TORONTO GOVERNING COUNCIL.

PI Fraser PE, Rommens JM, St George-Hyslop PH;

XX WPI; 1996-497631/49.

PT New presenilin genes - useful for diagnosis, therapy and drug
 PT screening of familial Alzheimer's disease, cerebral disorders, etc.

XX Claim 3; Page -; 178pp; English.

CC AAW05736-W05760 represent mutated versions of the human presenilin-1-1
 CC protein (see AAW05733 for wild type sequence). AAW05734 represents a
 CC different wild type form of presenilin-1 that results from alternate
 CC splicing of the genomic DNA sequence. The presenilins are a family of
 CC highly conserved integral membrane proteins with a common structural
 CC motif, common alternate splicing patterns, and common mutational hot
 CC spot regions. Mutations in PS genes are implicated in familial
 CC Alzheimer's disease (AD) and possibly other diseases such as cerebral
 CC haemorrhage, schizophrenia, depression etc., so detection of mutations in
 CC the DNA encoding the wild type sequences can be used for diagnosis of
 CC these diseases. The wild type proteins, or vectors that express them or
 CC containing antisense sequences, antibodies selective for these mutant
 CC forms of the proteins and modulators of PS gene expression are
 CC potentially useful for treatment of AD etc. Transgenic animals are

CC useful as models for drug screening. The antibodies can also be used e.g.
 CC for affinity purification and in immunoassays.

XX Sequence 467 AA;

Query Match 100.0%; Score 15; DB 17; Length 467;
 Best Local Similarity 100.0%; Pred. No. 3.6e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPHRSTPESRAA 15
 |||||
 DB 346 SHLGPHRSTPESRAA 360

RESULT 22

AAW05754 ID AAW05754 standard; Protein; 467 AA.

XX AAW05754;

DT 23-JUL-1997 (first entry)

DE Presenilin-1-1 A285V mutation.

KW Presenilin-1; human; hps1-1; hps1-2; PS-2; Integral membrane protein; AD;
 KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
 KW depression; antibody; gene expression modulator; therapy; mutin.

OS Homo sapiens.

FH Key Location/Qualifiers
 FT Modified-site 285
 FT /label= A285V

PN WO9634099-A2.

XX 31-OCT-1996.

PF 29-APR-1996; 96WO-CA00263.

XX 31-JUL-1995; 95US-0509359.

PR 28-APR-1995; 95US-0431048.

XX 28-JUN-1995; 95US-0496841.

PA (HSCR-) HSC RES & DEV LP.
 PA (UTOR) UNIV TORONTO GOVERNING COUNCIL.

PI Fraser PE, Rommens JM, St George-Hyslop PH;

XX WPI; 1996-497631/49.

PT New presenilin genes - useful for diagnosis, therapy and drug
 PT screening of familial Alzheimer's disease, cerebral disorders, etc.

XX Claim 3; Page -; 178pp; English.

CC AAW05736-W05760 represent mutated versions of the human presenilin-1-1
 CC protein (see AAW05733 for wild type sequence). AAW05734 represents a
 CC different wild type form of presenilin-1 that results from alternate
 CC splicing of the genomic DNA sequence. The presenilins are a family of
 CC highly conserved integral membrane proteins with a common structural
 CC motif, common alternate splicing patterns, and common mutational hot
 CC spot regions. Mutations in PS genes are implicated in familial
 CC Alzheimer's disease (AD) and possibly other diseases such as cerebral
 CC haemorrhage, schizophrenia, depression etc., so detection of mutations in
 CC the DNA encoding the wild type sequences can be used for diagnosis of
 CC these diseases. The wild type proteins, or vectors that express them or
 CC containing antisense sequences, antibodies selective for these mutant
 CC forms of the proteins and modulators of PS gene expression are
 CC potentially useful for treatment of AD etc. Transgenic animals are
 CC useful as models for drug screening. The antibodies can also be used e.g.
 CC for affinity purification and in immunoassays.

SO Sequence 467 AA;
Query Match 100.0%; Score 15; DB 17; Length 467;
Best Local Similarity 100.0%; Pred. No. 3.6e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLPHRSTPESRAA 15
|||||
DB 346 SHLPHRSTPESRAA 360

RESULT 23
AAW05755
ID AAW05755 standard; Protein; 467 AA.
XX
AC AAW05755;
XX
DT 23-JUL-1997 (first entry)
XX
DE Presenilin-1 L286V mutation.
XX
KW Presenilin-1; human; hPS1-1; hPS1-2; PS-2; Integral membrane protein; AD;
KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
KW depression; antibody; gene expression modulator; therapy; mutain.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 286
FT /Label= L286V
XX
PN WO9634099-A2.
XX
PD 31-OCT-1996.
XX
PF 29-APR-1996; 96MO-CA00263.
XX
PR 31-JUL-1995; 95US-0509359.
PR 28-APR-1995; 95US-0431048.
PR 28-JUN-1995; 95US-0496841.
XX
PA (HSCR-) HSC RES & DEV LP.
PA (UTOR) UNIV TORONTO GOVERNING COUNCIL.
XX
PI Fraser PE, Rommens JM, St George-Hyslop PH;
XX
DR WPI: 1996-497631/49.
XX
PT New presenilin genes - useful for diagnosis, therapy and drug
PT screening of familial Alzheimer's disease, cerebral disorders, etc.
XX
PS Claim 3; Page -: 178pp; English.
XX
CC AAW05736-W05760 represent mutated versions of the human presenilin-1-1
CC protein (see AAW05733 for wild type sequence). AAW05734 represents a
CC different wild type form of presenilin-1 that results from alternate
CC splicing of the genomic DNA sequence. The presenilins are a family of
CC highly conserved integral membrane proteins with a common structural
CC motif, common alternate splicing patterns, and common mutational hot
CC spot regions. Mutations in PS genes are implicated in familial
CC Alzheimer's disease (AD) and possibly other diseases such as cerebral
CC haemorrhage, schizophrenia, depression etc., so detection of mutations in
CC the DNA encoding the wild type sequences can be used for diagnosis of
CC these diseases. The wild type proteins, or vectors that express them or
CC containing antisense sequences, antibodies selective for these mutant
CC forms of the proteins and modulators of PS gene expression are
CC potentially useful for treatment of AD etc. Transgenic animals are
CC useful as models for drug screening. The antibodies can also be used e.g.
CC for affinity purification and in immunoassays.
XX
SQ Sequence 467 AA;

Query Match 100.0%; Score 15; DB 17; Length 467;

Best Local Similarity 100.0%; Pred. No. 3.6e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLPHRSTPESRAA 15
|||||
DB 346 SHLPHRSTPESRAA 360

RESULT 24
AAW05757
ID AAW05757 standard; Protein; 467 AA.
XX
AC AAW05757;
XX
DT 23-JUL-1997 (first entry)
XX
DE Presenilin-1 G384A mutation.
XX
KW Presenilin-1; human; hPS1-1; hPS1-2; PS-2; Integral membrane protein; AD;
KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
KW depression; antibody; gene expression modulator; therapy; mutain.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 384
FT /Label= G384A
XX
PN WO9634099-A2.
XX
PD 31-OCT-1996.
XX
PF 29-APR-1996; 96MO-CA00263.
XX
PR 31-JUL-1995; 95US-0509359.
PR 28-APR-1995; 95US-0431048.
PR 28-JUN-1995; 95US-0496841.
XX
PA (HSCR-) HSC RES & DEV LP.
PA (UTOR) UNIV TORONTO GOVERNING COUNCIL.
XX
PI Fraser PE, Rommens JM, St George-Hyslop PH;
XX
DR WPI: 1996-497631/49.
XX
PT New presenilin genes - useful for diagnosis, therapy and drug
PT screening of familial Alzheimer's disease, cerebral disorders, etc.
XX
PS Claim 3; Page -: 178pp; English.
XX
CC AAW05736-W05760 represent mutated versions of the human presenilin-1-1
CC protein (see AAW05733 for wild type sequence). AAW05734 represents a
CC different wild type form of presenilin-1 that results from alternate
CC splicing of the genomic DNA sequence. The presenilins are a family of
CC highly conserved integral membrane proteins with a common structural
CC motif, common alternate splicing patterns, and common mutational hot
CC spot regions. Mutations in PS genes are implicated in familial
CC Alzheimer's disease (AD) and possibly other diseases such as cerebral
CC haemorrhage, schizophrenia, depression etc., so detection of mutations in
CC the DNA encoding the wild type sequences can be used for diagnosis of
CC these diseases. The wild type proteins, or vectors that express them or
CC containing antisense sequences, antibodies selective for these mutant
CC forms of the proteins and modulators of PS gene expression are
CC potentially useful for treatment of AD etc. Transgenic animals are
CC useful as models for drug screening. The antibodies can also be used e.g.
CC for affinity purification and in immunoassays.
XX
SQ Sequence 467 AA;

Query Match 100.0%; Score 15; DB 17; Length 467;
Best Local Similarity 100.0%; Pred. No. 3.6e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

AAW05733
 ID AAW05733 standard; Protein: 467 AA.
 AC AAW05733:
 XX 23-JUL-1997 (first entry)
 DT
 XX
 DE Presenilin-1.
 XX
 KW Presenilin-1; human; hps1-1; hps1-2; PS-2; Integral membrane protein; AD;
 KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
 KW depression; antibody; gene expression modulator; therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO9604099-A2.
 XX
 PD 31-OCT-1996.
 XX
 PF 29-APR-1996; 96WO-CA00263.
 XX
 PR 31-JUL-1996; 95US-0509359.
 PR 28-APR-1995; 95US-0431048.
 PR 28-JUN-1995; 95US-0496841.
 XX
 PA (HSCR-) HSC RES & DEV LP.
 PA (UTOR) UNIV TORONTO GOVERNING COUNCIL.
 PI Fraser PE, Rommens JM, St George-Hyslop PH;
 DR WPI: 1996-497631/49.
 DR N-PSDB; AAT40028.
 XX
 PT New presenilin genes - useful for diagnosis, therapy and drug
 PT screening of familial Alzheimer's disease, cerebral disorders, etc.
 XX
 XX Claim 2; Page 128-130; 178pp; English.
 CC AAW05733 and AAW05734 represent the two different forms of wild type
 CC human presenilin-1 (PS-1). The form represented by AAW05734 results from
 CC alternate splicing of the genomic DNA sequence. AAW05762 represents the
 CC coding sequence for wild type human PS-2. The presenilins are a family
 CC of highly conserved integral membrane proteins with a common structural
 CC motif, common alternate splicing patterns, and common mutational hot
 CC spot regions. Mutations in PS genes are implicated in familial
 CC Alzheimer's disease (AD) and possibly other diseases such as cerebral
 CC haemorrhage, schizophrenia, depression etc., so detection of mutations in
 CC the DNA encoding these sequences can be used for diagnosis of these
 CC diseases. These proteins, or vectors that express them or containing
 CC antisense sequences, antibodies selective for mutant forms of these
 CC proteins (such as AAW05736) and modulators of PS gene expression are
 CC potentially useful for treatment of AD etc. Transgenic animals are useful
 CC as models for drug screening. The antibodies can also be used e.g. for
 CC affinity purification and in immunoassays.
 CC
 XX
 SO Sequence 467 AA;
 QY
 Query Match 100.0%; Score 15; DB 17; Length 467;
 Best Local Similarity 100.0%; Pred. No. 3.6e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1 SHLGHRSSTPESRAA 15
 346 SHLGHRSSTPESRAA 360
 RESULT 28
 AAW05735
 ID AAW05735 standard; Protein: 467 AA.
 AC AAW05735:
 XX 23-JUL-1997 (first entry)
 DT
 XX

XX
 DE Murine presenilin.
 XX
 KW Presenilin-1; mouse; hps1-1; hps1-2; PS-2; Integral membrane protein; AD;
 KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
 KW depression; antibody; gene expression modulator; therapy.
 XX
 OS Mus musculus.
 XX
 PN WO9604099-A2.
 XX
 PD 31-OCT-1996.
 XX
 PF 29-APR-1996; 96WO-CA00263.
 XX
 PR 31-JUL-1996; 95US-0509359.
 PR 28-APR-1995; 95US-0431048.
 PR 28-JUN-1995; 95US-0496841.
 XX
 PA (HSCR-) HSC RES & DEV LP.
 PA (UTOR) UNIV TORONTO GOVERNING COUNCIL.
 PI Fraser PE, Rommens JM, St George-Hyslop PH;
 DR WPI: 1996-497631/49.
 DR N-PSDB; AAT40030.
 XX
 PT New presenilin genes - useful for diagnosis, therapy and drug
 PT screening of familial Alzheimer's disease, cerebral disorders, etc.
 XX
 XX Claim 2; Page 145-146; 178pp; English.
 CC This sequence represents the wild type murine presenilin-1.
 CC AAW05733 and AAW05734 represent the two different forms of wild type
 CC human presenilin-1 (PS-1). The form represented by AAW05734 results from
 CC alternate splicing of the genomic DNA sequence. AAW05762 represents the
 CC coding sequence for wild type human PS-2. The presenilins are a family
 CC of highly conserved integral membrane proteins with a common structural
 CC motif, common alternate splicing patterns, and common mutational hot
 CC spot regions. Mutations in PS genes are implicated in familial
 CC Alzheimer's disease (AD) and possibly other diseases such as cerebral
 CC haemorrhage, schizophrenia, depression etc., so detection of mutations in
 CC the DNA encoding these sequences can be used for diagnosis of these
 CC diseases. These proteins, or vectors that express them or containing
 CC antisense sequences, antibodies selective for mutant forms of these
 CC proteins (such as AAW05736) and modulators of PS gene expression are
 CC potentially useful for treatment of AD etc. Transgenic animals are useful
 CC as models for drug screening. The antibodies can also be used e.g. for
 CC affinity purification and in immunoassays.
 CC
 XX
 SO Sequence 467 AA;
 QY
 Query Match 100.0%; Score 15; DB 17; Length 467;
 Best Local Similarity 100.0%; Pred. No. 3.6e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1 SHLGHRSSTPESRAA 15
 346 SHLGHRSSTPESRAA 360
 RESULT 29
 AAW05736
 ID AAW05736 standard; Protein: 467 AA.
 AC AAW05736:
 XX 23-JUL-1997 (first entry)
 DT
 XX
 DE Presenilin-1-1 A79X mutation.
 XX
 KW Presenilin-1; human; hps1-1; hps1-2; PS-2; Integral membrane protein; AD;
 KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;

KM depression; antibody; gene expression modulator; therapy; mutein.
 XX Homo sapiens.
 OS
 XX
 XX Key Location/Qualifiers
 FT Modified-site 79 /label= A79X
 FT /note= "x = unspecified amino acid"
 FT
 FT
 PN MO9634099-A2.
 XX
 XX 31-OCT-1996.
 PD
 XX
 XX 29-APR-1996; 96MO-CA00263.
 PF
 XX 31-JUL-1995; 95US-0509359.
 PR 28-APR-1995; 95US-0431048.
 PR 28-JUN-1995; 95US-0496841.
 XX
 XX (HSCR-) HSC RES & DEV LP.
 PA (UTOR) UNIV TORONTO GOVERNING COUNCIL.
 XX
 XX Fraser PE, Rommens JM, St George-Hyslop PH;
 PI WPI; 1996-497631/49.
 DR
 XX
 XX New presenilin genes - useful for diagnosis, therapy and drug
 PT screening of familial Alzheimer's disease, cerebral disorders, etc.
 XX
 XX Claim 3; Page -: 178pp; English.
 PS
 XX AAM05736-W05760 represent mutated versions of the human presenilin-1-1
 CC protein (see AAM05733 for wild type sequence). AAM05734 represents a
 CC different wild type form of presenilin-1 that results from alternate
 CC splicing of the genomic DNA sequence. The presenilins are a family of
 CC highly conserved integral membrane proteins with a common structural
 CC motif, common alternate splicing patterns, and common mutational hot
 CC spot regions. Mutations in PS genes are implicated in familial
 CC Alzheimer's disease (AD) and possibly other diseases such as cerebral
 CC haemorrhage, schizophrenia, depression etc., so detection of mutations in
 CC the DNA encoding the wild type sequences can be used for diagnosis of
 CC these diseases. The wild type proteins, or vectors that express them or
 CC containing antisense sequences, antibodies selective for these mutant
 CC forms of the proteins and modulators of PS gene expression are
 CC potentially useful for treatment of AD etc. Transgenic animals are
 CC useful as models for drug screening. The antibodies can also be used e.g.
 CC for affinity purification and in immunoassays.
 CC
 XX
 XX Sequence 467 AA:
 SO
 Query Match 100.0%; Score 15; DB 17; Length 467;
 Best Local Similarity 100.0%; Pred. No. 3.6e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 SHLGPHRSTPESRAA 15
 DB 346 SHLGPHRSTPESRAA 360
 RESULT 30
 AAM05737
 ID AAM05737 standard; Protein; 467 AA.
 XX
 XX AAM05737;
 AC
 XX
 XX 23-JUL-1997 (first entry)
 DT
 XX Presentin-1-1 V82L mutation.
 DE
 XX Presentin-1; human; hpsl-1; hpsl-2; PS-2; integral membrane protein; AD;
 KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
 KW depression; antibody; gene expression modulator; therapy; mutein.
 XX

OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT Modified-site 82 /label= V82L
 FT
 FT
 PN MO9634099-A2.
 XX
 XX 31-OCT-1996.
 PD
 XX
 XX 29-APR-1996; 96MO-CA00263.
 PF
 XX 31-JUL-1995; 95US-0509359.
 PR 28-APR-1995; 95US-0431048.
 PR 28-JUN-1995; 95US-0496841.
 XX
 XX (HSCR-) HSC RES & DEV LP.
 PA (UTOR) UNIV TORONTO GOVERNING COUNCIL.
 XX
 XX Fraser PE, Rommens JM, St George-Hyslop PH;
 PI WPI; 1996-497631/49.
 DR
 XX
 XX New presenilin genes - useful for diagnosis, therapy and drug
 PT screening of familial Alzheimer's disease, cerebral disorders, etc.
 XX
 XX Claim 3; Page -: 178pp; English.
 PS
 XX AAM05736-W05760 represent mutated versions of the human presenilin-1-1
 CC protein (see AAM05733 for wild type sequence). AAM05734 represents a
 CC different wild type form of presenilin-1 that results from alternate
 CC splicing of the genomic DNA sequence. The presenilins are a family of
 CC highly conserved integral membrane proteins with a common structural
 CC motif, common alternate splicing patterns, and common mutational hot
 CC spot regions. Mutations in PS genes are implicated in familial
 CC Alzheimer's disease (AD) and possibly other diseases such as cerebral
 CC haemorrhage, schizophrenia, depression etc., so detection of mutations in
 CC the DNA encoding the wild type sequences can be used for diagnosis of
 CC these diseases. The wild type proteins, or vectors that express them or
 CC containing antisense sequences, antibodies selective for these mutant
 CC forms of the proteins and modulators of PS gene expression are
 CC potentially useful for treatment of AD etc. Transgenic animals are
 CC useful as models for drug screening. The antibodies can also be used e.g.
 CC for affinity purification and in immunoassays.
 CC
 XX
 XX Sequence 467 AA:
 SO
 Query Match 100.0%; Score 15; DB 17; Length 467;
 Best Local Similarity 100.0%; Pred. No. 3.6e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 SHLGPHRSTPESRAA 15
 DB 346 SHLGPHRSTPESRAA 360
 RESULT 31
 AAM05738
 ID AAM05738 standard; Protein; 467 AA.
 XX
 XX AAM05738;
 AC
 XX
 XX 23-JUL-1997 (first entry)
 DT
 XX Presentin-1-1 V96F mutation.
 DE
 XX Presentin-1; human; hpsl-1; hpsl-2; PS-2; integral membrane protein; AD;
 KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
 KW depression; antibody; gene expression modulator; therapy; mutein.
 XX
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH

FT Modified-site 96
XX /label= v96f
PN W09634099-A2.
XX 31-OCT-1996.
PD 31-OCT-1996.
XX
PE 29-APR-1996; 96WO-CA00263.
XX
PR 31-JUL-1995; 95US-0509359.
PR 28-APR-1995; 95US-0431048.
PR 28-JUN-1995; 95US-0496841.
XX
PA (HSCR-) HSC RES & DEV LP.
PA (UTOR) UNIV TORONTO GOVERNING COUNCIL.
PI Fraser PE, Rommens JM, St George-Hyslop PH;
PI WPI; 1996-497631/49.
DR
XX
XX New presenilin genes - useful for diagnosis, therapy and drug
PT screening of familial Alzheimer's disease, cerebral disorders, etc.
XX
XX Claim 3; Page -: 178pp; English.
XX
CC AAM05736-W05760 represent mutated versions of the human presenilin-1-1
CC protein (see AAM05733 for wild type sequence). AAM05734 represents a
CC different wild type form of presenilin-1 that results from alternate
CC splicing of the genomic DNA sequence. The presenilins are a family of
CC highly conserved integral membrane proteins with a common structural
CC motif, common alternate splicing patterns, and common mutational hot
CC spot regions. Mutations in PS genes are implicated in familial
CC Alzheimer's disease (AD) and possibly other diseases such as cerebral
CC hemorrhage, schizophrenia, depression etc., so detection of mutations in
CC the DNA encoding the wild type sequences can be used for diagnosis of
CC these diseases. The wild type proteins, or vectors that express them or
CC containing antisense sequences, antibodies selective for these mutant
CC forms of the proteins and modulators of PS gene expression are
CC potentially useful for treatment of AD etc. Transgenic animals are
CC useful as models for drug screening. The antibodies can also be used e.g.
CC for affinity purification and in immunoassays.
XX
SQ Sequence 467 AA;
XX
Query Match 100.0%; Score 15; DB 17; Length 467;
Best Local Similarity 100.0%; Pred. No. 3.6e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHLGPHRSTPESRAA 15
DB 346 SHLGPHRSTPESRAA 360
RESULT 32
AAM05739
ID AAM05739 standard; Protein: 467 AA.
XX
AC AAM05739;
XX
DT 23-JUL-1997 (first entry)
XX
DE Presenilin-1-1 Y115H mutation.
XX
KW Presenilin-1; human; hPS1-1; hPS1-2; PS-2; Integral membrane protein; AD;
KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
KW depression; antibody; gene expression modulator; therapy; mutain.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 115
FT /label= y115h
XX

PN W09634099-A2.
XX 31-OCT-1996.
PD 31-OCT-1996.
XX
PE 29-APR-1996; 96WO-CA00263.
XX
PR 31-JUL-1995; 95US-0509359.
PR 28-APR-1995; 95US-0431048.
PR 28-JUN-1995; 95US-0496841.
XX
PA (HSCR-) HSC RES & DEV LP.
PA (UTOR) UNIV TORONTO GOVERNING COUNCIL.
PI Fraser PE, Rommens JM, St George-Hyslop PH;
PI WPI; 1996-497631/49.
DR
XX
XX New presenilin genes - useful for diagnosis, therapy and drug
PT screening of familial Alzheimer's disease, cerebral disorders, etc.
XX
XX Claim 3; Page -: 178pp; English.
XX
CC AAM05736-W05760 represent mutated versions of the human presenilin-1-1
CC protein (see AAM05733 for wild type sequence). AAM05734 represents a
CC different wild type form of presenilin-1 that results from alternate
CC splicing of the genomic DNA sequence. The presenilins are a family of
CC highly conserved integral membrane proteins with a common structural
CC motif, common alternate splicing patterns, and common mutational hot
CC spot regions. Mutations in PS genes are implicated in familial
CC Alzheimer's disease (AD) and possibly other diseases such as cerebral
CC hemorrhage, schizophrenia, depression etc., so detection of mutations in
CC the DNA encoding the wild type sequences can be used for diagnosis of
CC these diseases. The wild type proteins, or vectors that express them or
CC containing antisense sequences, antibodies selective for these mutant
CC forms of the proteins and modulators of PS gene expression are
CC potentially useful for treatment of AD etc. Transgenic animals are
CC useful as models for drug screening. The antibodies can also be used e.g.
CC for affinity purification and in immunoassays.
XX
SQ Sequence 467 AA;
XX
Query Match 100.0%; Score 15; DB 17; Length 467;
Best Local Similarity 100.0%; Pred. No. 3.6e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHLGPHRSTPESRAA 15
DB 346 SHLGPHRSTPESRAA 360
RESULT 33
AAM05740
ID AAM05740 standard; Protein: 467 AA.
XX
AC AAM05740;
XX
DT 23-JUL-1997 (first entry)
XX
DE Presenilin-1-1 M139V mutation.
XX
KW Presenilin-1; human; hPS1-1; hPS1-2; PS-2; Integral membrane protein; AD;
KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
KW depression; antibody; gene expression modulator; therapy; mutain.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 139
FT /label= M139X
FT /note= "X" - Thr, Val"
PN W09634099-A2.
XX

PD 31-OCT-1996.
 XX 29-APR-1996: 96WO-CA00263.
 XX 31-JUL-1995: 95US-0509359.
 PR 28-APR-1995: 95US-0431048.
 PR 28-JUN-1995: 95US-0496841.
 XX (HSCR-) HSC RES & DEV LP.
 PA (UTOR) UNIV TORONTO GOVERNING COUNCIL.
 XX Fraser PE, Rommens JM, St George-Hyslop PH;
 PI WPI; 1996-497631/49.
 DR WPI; 1996-497631/49.
 XX New presenilin genes - useful for diagnosis, therapy and drug
 PT screening of familial Alzheimer's disease, cerebral disorders, etc.
 XX Claim 3; Page -: 178pp; English.
 PS AAM05736-W05760 represent mutated versions of the human presenilin-1-1
 CC protein (see AAM05733 for wild type sequence). AAM05734 represents a
 CC different wild type form of presenilin-1 that results from alternate
 CC splicing of the genomic DNA sequence. The presenilins are a family of
 CC highly conserved integral membrane proteins with a common structural
 CC motif, common alternate splicing patterns, and common mutational hot
 CC spot regions. Mutations in PS genes are implicated in familial
 CC Alzheimer's disease (AD) and possibly other diseases such as cerebral
 CC haemorrhage, schizophrenia, depression etc., so detection of mutations in
 CC the DNA encoding the wild type sequences can be used for diagnosis of
 CC these diseases. The wild type proteins, or vectors that express them or
 CC containing antisense sequences, antibodies selective for these mutant
 CC forms of the proteins and modulators of PS gene expression are
 CC potentially useful for treatment of AD etc. Transgenic animals are
 CC useful as models for drug screening. The antibodies can also be used e.g.
 CC for affinity purification and in immunoassays.
 XX Sequence 467 AA:
 SQ
 Query Match 100.0%; Score 15; DB 17; Length 467;
 Best Local Similarity 100.0%; Pred. No. 3.6e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SHLGPHRSTPESRAA 15
 Db 346 SHLGPHRSTPESRAA 360
 RESULT 34
 AAM05741
 ID AAM05741 standard; Protein; 467 AA.
 AC AAM05741;
 XX 23-JUL-1997 (first entry)
 DT Presentin-1-1 I143T mutation.
 DE Presentin-1-1 I143T mutation.
 XX Presenilin-1; human; hpsl-1; hpsl-2; PS-2; integral membrane protein; AD;
 KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
 KW depression; antibody; gene expression modulator; therapy; mutein.
 XX Homo sapiens.
 OS Homo sapiens.
 XX Key Location/Qualifiers
 FH Modified-site 143
 FT /label= I143T
 FT
 XX WO9634099-A2.
 PN 31-OCT-1996.
 PD 31-OCT-1996.
 XX 29-APR-1996: 96WO-CA00263.

XX 31-JUL-1995: 95US-0509359.
 PR 28-APR-1995: 95US-0431048.
 PR 28-JUN-1995: 95US-0496841.
 XX (HSCR-) HSC RES & DEV LP.
 PA (UTOR) UNIV TORONTO GOVERNING COUNCIL.
 XX Fraser PE, Rommens JM, St George-Hyslop PH;
 PI WPI; 1996-497631/49.
 DR WPI; 1996-497631/49.
 XX New presenilin genes - useful for diagnosis, therapy and drug
 PT screening of familial Alzheimer's disease, cerebral disorders, etc.
 XX Claim 3; Page -: 178pp; English.
 PS AAM05736-W05760 represent mutated versions of the human presenilin-1-1
 CC protein (see AAM05733 for wild type sequence). AAM05734 represents a
 CC different wild type form of presenilin-1 that results from alternate
 CC splicing of the genomic DNA sequence. The presenilins are a family of
 CC highly conserved integral membrane proteins with a common structural
 CC motif, common alternate splicing patterns, and common mutational hot
 CC spot regions. Mutations in PS genes are implicated in familial
 CC Alzheimer's disease (AD) and possibly other diseases such as cerebral
 CC haemorrhage, schizophrenia, depression etc., so detection of mutations in
 CC the DNA encoding the wild type sequences can be used for diagnosis of
 CC these diseases. The wild type proteins, or vectors that express them or
 CC containing antisense sequences, antibodies selective for these mutant
 CC forms of the proteins and modulators of PS gene expression are
 CC potentially useful for treatment of AD etc. Transgenic animals are
 CC useful as models for drug screening. The antibodies can also be used e.g.
 CC for affinity purification and in immunoassays.
 XX Sequence 467 AA:
 SQ
 Query Match 100.0%; Score 15; DB 17; Length 467;
 Best Local Similarity 100.0%; Pred. No. 3.6e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SHLGPHRSTPESRAA 15
 Db 346 SHLGPHRSTPESRAA 360
 RESULT 35
 AAM05742
 ID AAM05742 standard; Protein; 467 AA.
 AC AAM05742;
 XX 23-JUL-1997 (first entry)
 DT Presentin-1-1 M146L/V mutation.
 DE Presentin-1-1 M146L/V mutation.
 XX Presenilin-1; human; hpsl-1; hpsl-2; PS-2; integral membrane protein; AD;
 KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
 KW depression; antibody; gene expression modulator; therapy; mutein.
 XX Homo sapiens.
 OS Homo sapiens.
 XX Key Location/Qualifiers
 FH Modified-site 146
 FT /label= M146X
 FT /note= "X = Leu, Val"
 FT
 XX WO9634099-A2.
 PN 31-OCT-1996.
 PD 31-OCT-1996.
 XX 29-APR-1996: 96WO-CA00263.
 PF 31-JUL-1995: 95US-0509359.


```

XX PI Fraser PE, Rommens JM, St George-Hyslop PH;
XX DR WPI: 1996-497631/49.
XX PT New presenilin genes - useful for diagnosis, therapy and drug
XX PT screening of familial Alzheimer's disease, cerebral disorders, etc.
XX PS Claim 3; Page -: 178pp; English.
XX CC AAM05736-W05760 represent mutated versions of the human presenilin-1-1
XX CC protein (see AAM05733 for wild type sequence). AAM05734 represents a
XX CC different wild type form of presenilin-1 that results from alternate
XX CC splicing of the genomic DNA sequence. The presenilins are a family of
XX CC highly conserved integral membrane proteins with a common structural
XX CC motif, common alternate splicing patterns, and common mutational hot
XX CC spot regions. Mutations in PS genes are implicated in familial
XX CC Alzheimer's disease (AD) and possibly other diseases such as cerebral
XX CC haemorrhage, schizophrenia, depression etc., so detection of mutations in
XX CC the DNA encoding the wild type sequences can be used for diagnosis of
XX CC these diseases. The wild type proteins, or vectors that express them or
XX CC containing antisense sequences, antibodies selective for these mutant
XX CC forms of the proteins and modulators of PS gene expression are
XX CC potentially useful for treatment of AD etc. Transgenic animals are
XX CC useful as models for drug screening. The antibodies can also be used e.g.
XX CC for affinity purification and in immunoassays.
XX SQ Sequence 467 AA;

Query Match 100.0%; Score 15; DB 17; Length 467;
Best Local Similarity 100.0%; Pred. No. 3.6e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGPBHRSTPESRAA 15
Db 346 SHLGPBHRSTPESRAA 360
|||||
|||||

RESULT 38
AAM05745 ID AAM05745 standard; Protein; 467 AA.
XX AC AAM05745;
XX DT 23-JUL-1997 (first entry)
XX DE Presenilin-1-1 G209V mutation.
XX KW Presenilin-1; human; hps1-1; hps1-2; PS-2; integral membrane protein; AD;
XX KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
XX KW depression; antibody; gene expression modulator; therapy; mutain.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Modified-site 209 /label= G209V
XX PN W09634099-A2.
XX PD 31-OCT-1996.
XX PF 29-APR-1996; 96WO-CA00263.
XX PR 31-JUL-1995; 95US-0509359.
XX PR 28-APR-1995; 95US-0431048.
XX PR 28-JUN-1995; 95US-0496841.
XX PA (HSCR-) HSC RES & DEV LP.
XX PA (UTOR ) UNIV TORONTO GOVERNING COUNCIL.
XX PI Fraser PE, Rommens JM, St George-Hyslop PH;
XX
```

```

DR WPI: 1996-497631/49.
XX PT New presenilin genes - useful for diagnosis, therapy and drug
XX PT screening of familial Alzheimer's disease, cerebral disorders, etc.
XX PS Claim 3; Page -: 178pp; English.
XX CC AAM05736-W05760 represent mutated versions of the human presenilin-1-1
XX CC protein (see AAM05733 for wild type sequence). AAM05734 represents a
XX CC different wild type form of presenilin-1 that results from alternate
XX CC splicing of the genomic DNA sequence. The presenilins are a family of
XX CC highly conserved integral membrane proteins with a common structural
XX CC motif, common alternate splicing patterns, and common mutational hot
XX CC spot regions. Mutations in PS genes are implicated in familial
XX CC Alzheimer's disease (AD) and possibly other diseases such as cerebral
XX CC haemorrhage, schizophrenia, depression etc., so detection of mutations in
XX CC the DNA encoding the wild type sequences can be used for diagnosis of
XX CC these diseases. The wild type proteins, or vectors that express them or
XX CC containing antisense sequences, antibodies selective for these mutant
XX CC forms of the proteins and modulators of PS gene expression are
XX CC potentially useful for treatment of AD etc. Transgenic animals are
XX CC useful as models for drug screening. The antibodies can also be used e.g.
XX CC for affinity purification and in immunoassays.
XX SQ Sequence 467 AA;

Query Match 100.0%; Score 15; DB 17; Length 467;
Best Local Similarity 100.0%; Pred. No. 3.6e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGPBHRSTPESRAA 15
Db 346 SHLGPBHRSTPESRAA 360
|||||
|||||

RESULT 39
AAM05746 ID AAM05746 standard; Protein; 467 AA.
XX AC AAM05746;
XX DT 23-JUL-1997 (first entry)
XX DE Presenilin-1-1 I211T mutation.
XX KW Presenilin-1; human; hps1-1; hps1-2; PS-2; integral membrane protein; AD;
XX KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
XX KW depression; antibody; gene expression modulator; therapy; mutain.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Modified-site 211 /label= I211T
XX PN W09634099-A2.
XX PD 31-OCT-1996.
XX PF 29-APR-1996; 96WO-CA00263.
XX PR 31-JUL-1995; 95US-0509359.
XX PR 28-APR-1995; 95US-0431048.
XX PR 28-JUN-1995; 95US-0496841.
XX PA (HSCR-) HSC RES & DEV LP.
XX PA (UTOR ) UNIV TORONTO GOVERNING COUNCIL.
XX PI Fraser PE, Rommens JM, St George-Hyslop PH;
XX DR WPI: 1996-497631/49.
XX PT New presenilin genes - useful for diagnosis, therapy and drug
```

PT screening of familial Alzheimer's disease, cerebral disorders, etc.
 XX
 PS Claim 3; Page -: 178pp; English.
 XX
 CC AAM05736-W05760 represent mutated versions of the human presenilin-1-1
 CC protein (see AAM05733 for wild type sequence). AAM05734 represents a
 CC different wild type form of presenilin-1 that results from alternate
 CC splicing of the genomic DNA sequence. The presenilins are a family of
 CC highly conserved integral membrane proteins with a common structural
 CC motif, common alternate splicing patterns, and common mutational hot
 CC spot regions. Mutations in PS genes are implicated in familial
 CC Alzheimer's disease (AD) and possibly other diseases such as cerebral
 CC haemorrhage, schizophrenia, depression etc., so detection of mutations in
 CC the DNA encoding the wild type sequences can be used for diagnosis of
 CC these diseases. The wild type proteins, or vectors that express them or
 CC containing antisense sequences, antibodies selective for these mutant
 CC forms of the proteins and modulators of PS gene expression are
 CC potentially useful for treatment of AD etc. Transgenic animals are
 CC useful as models for drug screening. The antibodies can also be used e.g.
 CC for affinity purification and in immunoassays.
 XX
 SQ Sequence 467 AA;
 XX
 Query Match 100.0%; Score 15; DB 17; Length 467;
 Best Local Similarity 100.0%; Pred. No. 3.6e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SHLGPHRSTPESRAA 15
 DB 346 SHLGPHRSTPESRAA 360
 |||||
 AAM05747 standard; Protein; 467 AA.
 XX
 AC AAM05747;
 XX
 DT 23-JUL-1997 (first entry)
 XX
 DE Presenilin-1-1 I231T mutation.
 XX
 KW Presenilin-1; human; hps1-1; hps1-2; PS-2; integral membrane protein; AD;
 KM familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
 KM depression; antibody; gene expression modulator; therapy; mutain.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 231
 FT /label= I231T
 XX
 PN WO9634099-A2.
 XX
 PD 31-OCT-1996.
 XX
 PE 29-APR-1996; 96WO-CA00263.
 XX
 PR 31-JUL-1995; 95US-0509359.
 PR 28-APR-1995; 95US-0431048.
 PR 28-JUN-1995; 95US-0496841.
 XX
 PA (HSCR-) HSC RES & DEV LP.
 PA (UTOR) UNIV TORONTO GOVERNING COUNCIL.
 PI Fraser PE, Rommens JM, St George-Hyslop PH;
 XX WPI: 1996-497631/49.
 DR
 PT New presenilin genes - useful for diagnosis, therapy and drug
 XX screening of familial Alzheimer's disease, cerebral disorders, etc.
 XX
 PS Claim 3; Page -: 178pp; English.

XX
 CC AAM05736-W05760 represent mutated versions of the human presenilin-1-1
 CC protein (see AAM05733 for wild type sequence). AAM05734 represents a
 CC different wild type form of presenilin-1 that results from alternate
 CC splicing of the genomic DNA sequence. The presenilins are a family of
 CC highly conserved integral membrane proteins with a common structural
 CC motif, common alternate splicing patterns, and common mutational hot
 CC spot regions. Mutations in PS genes are implicated in familial
 CC Alzheimer's disease (AD) and possibly other diseases such as cerebral
 CC haemorrhage, schizophrenia, depression etc., so detection of mutations in
 CC the DNA encoding the wild type sequences can be used for diagnosis of
 CC these diseases. The wild type proteins, or vectors that express them or
 CC containing antisense sequences, antibodies selective for these mutant
 CC forms of the proteins and modulators of PS gene expression are
 CC potentially useful for treatment of AD etc. Transgenic animals are
 CC useful as models for drug screening. The antibodies can also be used e.g.
 CC for affinity purification and in immunoassays.
 XX
 SQ Sequence 467 AA;
 XX
 Query Match 100.0%; Score 15; DB 17; Length 467;
 Best Local Similarity 100.0%; Pred. No. 3.6e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SHLGPHRSTPESRAA 15
 DB 346 SHLGPHRSTPESRAA 360
 |||||
 AAM05748 standard; Protein; 467 AA.
 XX
 AC AAM05748;
 XX
 DT 23-JUL-1997 (first entry)
 XX
 DE Presenilin-1-1 A246E mutation.
 XX
 KW Presenilin-1; human; hps1-1; hps1-2; PS-2; integral membrane protein; AD;
 KM familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
 KM depression; antibody; gene expression modulator; therapy; mutain.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 246
 FT /label= A246E
 XX
 PN WO9634099-A2.
 XX
 PD 31-OCT-1996.
 XX
 PE 29-APR-1996; 96WO-CA00263.
 XX
 PR 31-JUL-1995; 95US-0509359.
 PR 28-APR-1995; 95US-0431048.
 PR 28-JUN-1995; 95US-0496841.
 XX
 PA (HSCR-) HSC RES & DEV LP.
 PA (UTOR) UNIV TORONTO GOVERNING COUNCIL.
 PI Fraser PE, Rommens JM, St George-Hyslop PH;
 XX WPI: 1996-497631/49.
 DR
 PT New presenilin genes - useful for diagnosis, therapy and drug
 XX screening of familial Alzheimer's disease, cerebral disorders, etc.
 XX
 PS Claim 3; Page -: 178pp; English.
 XX
 CC AAM05736-W05760 represent mutated versions of the human presenilin-1-1
 CC protein (see AAM05733 for wild type sequence). AAM05734 represents a

CC different wild type form of presenilin-1 that results from alternate
 CC splicing of the genomic DNA sequence. The presenilins are a family of
 CC highly conserved integral membrane proteins with a common structural
 CC motif, common alternate splicing patterns, and common mutational hot
 CC spot regions. Mutations in PS genes are implicated in familial
 CC Alzheimer's disease (AD) and possibly other diseases such as cerebral
 CC haemorrhage, schizophrenia, depression etc., so detection of mutations in
 CC the DNA encoding the wild type sequences can be used for diagnosis of
 CC these diseases. The wild type proteins, or vectors that express them or
 CC containing antisense sequences, antibodies selective for these mutant
 CC forms of the proteins and modulators of PS gene expression are
 CC potentially useful for treatment of AD etc. Transgenic animals are
 CC useful as models for drug screening. The antibodies can also be used e.g.
 CC for affinity purification and in immunoassays.

CC Sequence 467 AA;

Query Match 100.0%; Score 15; DB 17; Length 467;

Best Local Similarity 100.0%; Pred. No. 3.6e-08; Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGPFRSTPESRAA 15

DB 346 SHLGPFRSTPESRAA 360

RESULT 42

AAW05749 standard; Protein; 467 AA.

AC AAW05749;

DT 23-JUL-1997 (first entry)

DE Presenilin-1-1 A260V mutation.

KW Presenilin-1; human; hpsl-1; hpsl-2; PS-2; integral membrane protein; AD;

KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;

KW depression; antibody; gene expression modulator; therapy; muteln.

OS Homo sapiens.

FT Modified-site 260

PN MO9634099-A2.

PD 31-OCT-1996.

PF 29-APR-1996; 96WO-CA00263.

PR 31-JUL-1995; 95US-0509359.

PR 28-APR-1995; 95US-0431048.

PR 28-JUN-1995; 95US-0496841.

PA (HSCR-1) HSC RES & DEV LP.

PI (UTOR) UNIV TORONTO GOVERNING CONCIL.

PI Fraser BE, Rommens JM, St George-Hyslop PH;

PT WPI; 1996-497631/49.

PS New presenilin genes - useful for diagnosis, therapy and drug

CC screening of familial Alzheimer's disease, cerebral disorders, etc.

CC Claim 3; Page -; 178pp; English.

CC AAW05736-W05760 represent mutated versions of the human presenilin-1-1

CC protein (see AAW05733 for wild type sequence). AAW05734 represents a

CC different wild type form of presenilin-1 that results from alternate

CC splicing of the genomic DNA sequence. The presenilins are a family of

CC highly conserved integral membrane proteins with a common structural

CC motif, common alternate splicing patterns, and common mutational hot
 CC spot regions. Mutations in PS genes are implicated in familial
 CC Alzheimer's disease (AD) and possibly other diseases such as cerebral
 CC haemorrhage, schizophrenia, depression etc., so detection of mutations in
 CC the DNA encoding the wild type sequences can be used for diagnosis of
 CC these diseases. The wild type proteins, or vectors that express them or
 CC containing antisense sequences, antibodies selective for these mutant
 CC forms of the proteins and modulators of PS gene expression are
 CC potentially useful for treatment of AD etc. Transgenic animals are
 CC useful as models for drug screening. The antibodies can also be used e.g.
 CC for affinity purification and in immunoassays.

CC Sequence 467 AA;

Query Match 100.0%; Score 15; DB 17; Length 467;

Best Local Similarity 100.0%; Pred. No. 3.6e-08; Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGPFRSTPESRAA 15

DB 346 SHLGPFRSTPESRAA 360

RESULT 43

AAW27177 standard; Protein; 467 AA.

AC AAW27177;

DT 09-DEC-1997 (first entry)

DE Human mutant S182 gene, PS1, product related to Alzheimer's disease.

KW Mutant; antisense; antibody; vaccine; Alzheimer's disease.

OS Homo sapiens.

FT WO9708319-A1.

PD 06-MAR-1997.

PF 03-SEP-1996; 96WO-US14114.

PR 30-AUG-1996; 96US-0706344.

PR 31-AUG-1995; 95US-0003054.

PA (GEHO) GEN HOSPITAL CORP.

PI Tanzi RE, Wasco W;

PI WPI; 1997-179276/16.

DR N-PSDB; AAT85333.

PT Chromosome 14 early-onset familial Alzheimer's disease gene PS1

PT mutants - useful for diagnosing likelihood of developing Alzheimer's

PT disease, also anti-sense sequences, antibodies and vaccines to delay

PT onset

PS Claim 13; Page 77-78; 99pp; English.

CC The present sequence represents the human mutant S182 gene, PS1,

CC product. Mutant PS1 produces a gene product that increases the

CC probability of Alzheimer's disease. A nucleic acid sequence able to

CC hybridise to sequences coding for a mutant PS1 polypeptide can be used

CC as probes for diagnosing an increased likelihood of contracting

CC Alzheimer's disease. Antibodies against the mutant polypeptide can also

CC be used for this purpose. Vectors containing or expressing a nucleic

CC acid molecule, protein or antibody specific for mutant PS1 can be

CC administered to a patient to reduce the likelihood, or delay the onset,

CC of Alzheimer's disease, e.g. anti-sense RNA expression can be used to

CC decrease expression of the PS1 peptide. Transgenic animals expressing

CC the Alzheimer's disease protein can be used to test candidate

CC therapeutics and to investigate the normal role of PS1. The PS1 peptide

CC may also be included in pharmaceutical compositions (vaccines) for
 CC Alzheimer's disease therapy.

XX
 SQ Sequence 467 AA;

Query Match 100.0%; Score 15; DB 18; Length 467;
 Best Local Similarity 100.0%; Pred. No. 3.6e-08;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPHRSTPESRAA 15
 |||||
 DB 346 SHLGPHRSTPESRAA 360

RESULT 44

AAW27176
 ID AAW27176 standard; Protein; 467 AA.

AC AAW27176;

DT 09-DEC-1997 (first entry)

DE Human S182 gene, P51 locus, product related to Alzheimer's disease.

XX Mutant; antisense; antibody; vaccine; Alzheimer's disease.

XX Homo sapiens.

XX W09708319-A1.

PD 06-MAR-1997.

PF 03-SEP-1996; 96WO-US14114.

PR 30-AUG-1996; 96US-0706344.

PR 31-AUG-1995; 95US-0003054.

PA (GEHO) GEN HOSPITAL CORP.

PI Tanzi RE, Wasco W;

DR WPI; 1997-179276/16.

DR N-PSDB; AAT85332.

PT Chromosome 14 early-onset familial Alzheimer's disease gene P51
 PT mutants - useful for diagnosing likelihood of developing Alzheimer's
 PT disease, also anti-sense sequences, antibodies and vaccines to delay
 PT onset

PS Claim 12; Page 72-73; 99pp; English.

CC The present sequence represents the human S182 gene, P51 locus, product.

CC Mutant P51 produces a gene product that increases the probability of
 CC Alzheimer's disease. A nucleic acid sequence able to hybridize to
 CC sequences coding for a mutant P51 polypeptide can be used as probes for
 CC diagnosing an increased likelihood of contracting Alzheimer's disease.

CC Antibodies against the mutant polypeptide can also be used for this
 CC purpose. Vectors containing or expressing a nucleic acid molecule,
 CC protein or antibody specific for mutant P51 can be administered to a
 CC patient to reduce the likelihood, or delay the onset, of Alzheimer's
 CC disease, e.g. anti-sense RNA expression can be used to decrease
 CC expression of the P51 peptide. Transgenic animals expressing the
 CC Alzheimer's disease protein can be used to test candidate therapeutics
 CC and to investigate the normal role of P51. The P51 peptide may also be
 CC included in pharmaceutical compositions (vaccines) for Alzheimer's
 CC disease therapy.

XX
 SQ Sequence 467 AA;

Query Match 100.0%; Score 15; DB 18; Length 467;
 Best Local Similarity 100.0%; Pred. No. 3.6e-08;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPHRSTPESRAA 15
 |||||
 DB 346 SHLGPHRSTPESRAA 360

RESULT 45

AAW11839
 ID AAW11839 standard; Protein; 467 AA.

AC AAW11839;

DT 07-MAY-1997 (first entry)

DE Human early onset Alzheimer's disease (EOAD) polypeptide.

XX Early onset Alzheimer's disease; EOAD; neurodegenerative disease;
 XX diagnosis; therapy; inhibitor; antagonist; antibody.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 26..29 /note="unidentified amino acid residues"

PN W09703086-A1.

PD 30-JAN-1997.

PF 26-JUN-1996; 96WO-US11064.

PR 18-JUL-1995; 95US-0001501.

PR 13-JUL-1995; 95US-0001142.

PA (UYSE-) UNIV SOUTH FLORIDA.

PI Hardy JA;

DR WPI; 1997-118980/11.

DR N-PSDB; AAT59535.

PT Early onset Alzheimer's disease gene - useful for diagnosing a
 PT pre-disposition to Alzheimer's disease

PS Disclosure; Fig 1; 44pp; English.

CC A 467-amino acid polypeptide (AAW11839) is the product of a
 CC full-length cDNA (AAT59535) of the early onset Alzheimer's disease
 CC (EOAD) gene. A 463-amino acid polypeptide (AAW11840) is the
 CC product of another full-length cDNA (AAT59536) of an EOAD splice
 CC variant gene. The 2 polypeptides can be produced in transformed
 CC host cells and used to raise antibodies, or to identify
 CC antagonist/inhibitor cpds., useful in the treatment of Alzheimer's
 CC disease, esp. EOAD.

XX
 SQ Sequence 467 AA;

Query Match 100.0%; Score 15; DB 18; Length 467;
 Best Local Similarity 100.0%; Pred. No. 3.6e-08;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHLGPHRSTPESRAA 15
 |||||
 DB 346 SHLGPHRSTPESRAA 360

RESULT 46

AAW56770
 ID AAW56770 standard; Protein; 467 AA.

AC AAW56770;

DT 13-OCT-1998 (first entry)

DE Homo sapiens PS-1.
XX
XX PS-1; presenilin; presenilin-1; PSP-1; Alzheimer's disease;
KW serine protease; neurodegeneration; predisposition; diagnosis.
XX
OS Homo sapiens.
PN EP828003-A2.
XX
XX 11-MAR-1998.
PD
XX
PF 26-AUG-1997; 97EP-0306501.
XX
PR 13-DEC-1996; 96US-0032875.
PR 06-SEP-1986; 96US-0025436.
PR 25-OCT-1996; 96US-0027873.
XX
PA (SMK) SMITHKLINE BECHAM CORP.
PA (SMK) SMITHKLINE BECHAM PLC.
XX
XX Browne MJ, Clinkenbeard HE, Creasy CL, Karran EH;
PI Livi GP, Southan CD;
XX
DR WPI; 1998-161101/15.
DR N-PSDB; AAV29525.
XX
XX Nucleic acids encoding human serum protease protein(s) - used for
PT diagnosing pre-disposition to Alzheimer's disease, etc.
XX
XX Example 1; Page 25-26; 65pp; English.
PS
XX The sequence is that of of presenilin PS-1 which was used
CC in the cloning and isolation of the serine protease PSP1.
CC
XX
SQ Sequence 467 AA;

Query Match 100.0%; Score 15; DB 19; Length 467;
Best Local Similarity 100.0%; Pred. No. 3.6e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGPHRSTPESRAA 15
DB 346 SHLGPHRSTPESRAA 360

RESULT 47
AAW23964
ID AAW23964 standard; protein; 467 AA.
XX
XX AAW23964;
AC
XX
XX 20-JUL-1998 (first entry)
DT
XX
XX Human presenilin-1.
DE
XX
XX Presenilin-1; PS1 gene; human; familial Alzheimer's disease; FAD;
KW cerebral haemorrhage; schizophrenia; depression; epilepsy;
KW mental retardation; diagnosis; therapy; transgenic animal.
XX
XX Homo sapiens.
OS
XX
XX
XX Key Location/Qualifiers
FH 82..100
FT /label= TM1
FT /note= "transmembrane domain 1"
FT 101..132
FT /label= TM1.2
FT /note= "hydrophilic loop"
FT 133..154
FT /label= TM2
FT /note= "transmembrane domain 2"
FT 155..163
FT /label= TM2.3
FT

FT /note= "hydrophilic loop"
FT 164..183
FT /label= TM3
FT /note= "transmembrane domain 3"
FT 184..194
FT /label= TM3.4
FT /note= "hydrophilic loop"
FT 195..212
FT /label= TM4
FT /note= "transmembrane domain 4"
FT 213..220
FT /label= TM4.5
FT /note= "hydrophilic loop"
FT 221..238
FT /label= TM5
FT /note= "transmembrane domain 5"
FT 239..243
FT /label= TM5.6
FT /note= "hydrophilic loop"
FT 244..262
FT /label= TM6
FT /note= "transmembrane domain 6"
FT 263..407
FT /label= TM6.7
FT /note= "hydrophilic loop"
FT 408..428
FT /label= TM8
FT /note= "transmembrane domain 8"
FT 177
FT /note= "Phe117Ser mutation site (Claim 1)"
FT 439
FT /note= "Ile439Val mutation site (Claim 1)"
FT 257
FT /note= "Asp257Ala mutation site, associated with
FT residue 258-290 deletion (Claim 1)"
FT 258..290
FT /note= "residue 258-290 deletion mutant, associated
FT with Asp257Ala mutation (Claim 1)"
FT 143
FT /note= "Ile143Thr mutation site (Claim 18)"
FT 146
FT /note= "Met146Leu mutation site (Claim 18)"
FT 171
FT /note= "Leu171Pro mutation site (Claim 18)"
FT 260
FT /note= "Ala260Val mutation site (Claim 18)"
FT 263
FT /note= "Cys263Arg mutation site (Claim 18)"
FT 264
FT /note= "Pro264Leu mutation site (Claim 18)"
FT 267
FT /note= "Pro267Ser mutation site (Claim 18)"
FT 280
FT /note= "Glu280Ala mutation site (Claim 18)"
FT 280
FT /note= "Glu280Gly mutation site (Claim 18)"
FT 285
FT /note= "Ala285Val mutation site (Claim 18)"
FT 286
FT /note= "Leu286Val mutation site (Claim 18)"
FT 322
FT /note= "Leu322Val mutation site (Claim 18)"
FT 392
FT /note= "Leu392Val mutation site (Claim 18)"
FT 410
FT /note= "Cys410Tyr mutation site (Claim 18)"
FT 79
FT /note= "Ala79Xaa mutation site"
FT 82
FT /note= "Val82Leu mutation site"
FT 96
FT /note= "Val96Phe mutation site"
FT 115
FT Misc-difference

FT	Misc-difference	139	/note-	"Tyr115His mutation site"
FT	Misc-difference	139	/note-	"Met139Thr mutation site"
FT	Misc-difference	139	/note-	"Met139Val mutation site"
FT	Misc-difference	146	/note-	"Met16Val mutation site"
FT	Misc-difference	163	/note-	"His163Arg mutation site"
FT	Misc-difference	163	/note-	"His163Tyr mutation site"
FT	Misc-difference	209	/note-	"Gly209Val mutation site"
FT	Misc-difference	211	/note-	"Ile211Thr mutation site"
FT	Misc-difference	231	/note-	"Ala231Thr mutation site"
FT	Misc-difference	246	/note-	"Ala246Glu mutation site"
FT	Misc-difference	291..319	/note-	"residue 291-319 deletion site"
FT	Misc-difference	384	/note-	"Gly384Ala mutation site"
FT	Misc-difference	26..29	/note-	"residue 26-29 deletion, resulting from alternative splicing"
PN	W0801549-A2.			
PD	15-JAN-1998.			
XX	04-JUL-1997;	97WO-CA00475.		
XX	02-JAN-1997;	97US-0034590.		
PR	05-JUL-1986;	96US-0021673.		
PR	12-JUL-1996;	96US-0021700.		
PR	08-NOV-1996	96US-0029895.		
PA	(HSCR-) HSC RES & DEV LP.			
PA	(UTOR) UNIV TORONTO GOVERNING COUNCIL.			
XX	Fraser PE, Rommens JM, St George-Hyslop PH;			
XX	WPI; 1998-286355/25.			
DR	N-PSDB; AAY04666.			
XX				
PT	New isolated mutant presenilin-1 genes - useful for developing			
PT	products for use in detection, diagnosis and therapy of Alzheimer's			
PT	disease and for drug screening			
XX				
PS	Claim 1; Page 180-182; 238pp; English.			
XX				
CC	This polypeptide comprises human presenilin-1 (hps1). Its amino			
CC	acid sequence was deduced from an isolated cDNA clone (see AAY04666)			
CC	Another hps1 sequence (see AAM23965) results from alternative			
CC	splicing of the hps1 mRNA transcript. A murine PS1 homologue (see			
CC	AAM23966) and a human presenilin-2 protein (see AAM23967) are also			
CC	provided. Mutations in the PS-1 and PS-2 genes are linked to the			
CC	development in humans of forms of familial Alzheimer's disease			
CC	(FAD) and may be causative of other disorders, e.g. cognitive,			
CC	intellectual, neurological or physiological disorders such as			
CC	cerebral haemorrhage, schizophrenia, depression, mental retardation			
CC	and epilepsy. Use of the nucleic acids and proteins comprising or			
CC	derived from the presenilins is made in screening and diagnosing			
CC	FAD, identifying and developing therapeutics for treatment of FAD,			
CC	and in producing cell lines and transgenic animals useful as models			
CC	of FAD. Methods for identifying substances that bind to, or			
CC	modulate the activity of a presenilin protein, and methods for			
CC	identifying substances that affect the interaction of a			
CC	presenilin-interacting protein with a presenilin protein are also			
CC	disclosed.			
XX				
XX	Sequence 467 AA;			
XQ				

Query Match	Similarity	100.0%	Score 15;	DB 19;	Length 467;
Best Local	Similarity	100.0%	Pred. No. 3.6e-08;		
Matches	15;	Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
Oy	1	SHLGPRTSPESRAA	15		
Db	346	SHLGPRTSPESRAA	360		
RESULT 48					
ID	AAW23966	AAW23966 standard; Protein; 467 AA.			
XX	AAW23966;				
AC	AAW23966;				
XX	20-JUL-1998	(first entry)			
DT					
XX					
DE	Mouse presenilin-1 homologue.				
XX					
KW	Presenilin-1; PS1 gene; mouse; familial Alzheimer's disease; FAD;				
KW	cerebral haemorrhage; schizophrenia; depression; epilepsy;				
KW	mental retardation; diagnosis; therapy; transgenic animal.				
XX					
OS	Mus musculus.				
XX					
FH	Key	Location/Qualifiers			
FH	Domain	82..100			
FT		/label= TM1			
FT		/note= "transmembrane domain 1"			
FT	Domain	101..132			
FT		/label= TM1-2			
FT		/note= "hydrophilic loop"			
FT	Domain	133..154			
FT		/label= TM2			
FT		/note= "transmembrane domain 2"			
FT	Domain	155..163			
FT		/label= TM2-3			
FT		/note= "hydrophilic loop"			
FT	Domain	164..183			
FT		/label= TM3			
FT		/note= "transmembrane domain 3"			
FT	Domain	184..194			
FT		/label= TM3-4			
FT		/note= "hydrophilic loop"			
FT	Domain	195..212			
FT		/label= TM4			
FT		/note= "transmembrane domain 4"			
FT	Domain	213..220			
FT		/label= TM4-5			
FT		/note= "hydrophilic loop"			
FT	Domain	221..238			
FT		/label= TM5			
FT		/note= "transmembrane domain 5"			
FT	Domain	239..243			
FT		/label= TM5-6			
FT		/note= "hydrophilic loop"			
FT	Domain	244..262			
FT		/label= TM6			
FT		/note= "transmembrane domain 6"			
FT	Domain	263..407			
FT		/label= TM6-7			
FT		/note= "hydrophilic loop"			
FT	Domain	408..428			
FT		/label= TM8			
FT		/note= "transmembrane domain 8"			
FT	Misc-difference	177			
FT		/note= "Phe17Ser mutation site (claim 1)"			
FT	Misc-difference	439			
FT		/note= "Ile439Val mutation site (claim 1)"			
XX					
PN	WO9801549-A2.				
XX					

PD 15-JAN-1998.
 XX
 PF 04-JUL-1997; 97WO-CA00475.
 XX
 PR 02-JAN-1997; 97US-0034590.
 PR 05-JUL-1996; 96US-0021673.
 PR 12-JUL-1996; 96US-0021700.
 PR 08-NOV-1996; 96US-0029895.
 XX
 PA (HSCR-) HSC RES & DEV LP.
 PA (UTOR) UNIV TORONTO GOVERNING COUNCIL.
 PI Fraser PE, Rommens JM, St George-Hyslop PH;
 XX
 DR WPI: 1998-286355/25.
 DR N-PSDB; AAV04668.
 XX
 PT New isolated mutant presenilin-1 genes - useful for developing
 PT products for use in detection, diagnosis and therapy of Alzheimer's
 PT disease and for drug screening
 XX
 PS Disclosure; page 199-200; 238pp; English.
 XX
 CC This polypeptide comprises the murine presenilin-1 (PS1) homologue.
 CC its amino acid sequence was deduced from an isolated cDNA clone
 CC (see AAV04668). Mutations in the human PS1 and PS2 genes (see
 CC AAV04666-68) have been linked to the development in humans of forms
 CC of familial Alzheimer's disease (FAD). All amino acids that are
 CC mutated in analysed FAD pedigrees (see AAW23964) were conserved in
 CC the murine homologue. Use of the nucleic acids and proteins
 CC comprising or derived from presenilins can be made in screening and
 CC diagnosing FAD, identifying and developing therapeutics for
 CC treatment of FAD, and in producing cell lines and transgenic
 CC animals useful as models of FAD. Methods for identifying
 CC substances that bind to, or modulate the activity of a presenilin
 CC protein, and methods for identifying substances that affect the
 CC interaction of a presenilin-interacting protein with a presenilin
 CC protein are also disclosed.
 XX
 SQ Sequence 467 AA;
 XX
 Query Match 100.0%; Score 15; DB 19; Length 467;
 Best Local Similarity 100.0%; Pred. No. 3.6e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SHIGPHRSTPESRAA 15
 |||||
 Db 346 SHIGPHRSTPESRAA 360

RESULT 49
 AAW41430
 ID AAW41430 standard; Protein; 467 AA.
 XX
 AC AAW41430;
 XX
 DT 04-JUN-1998 (first entry)
 XX
 DE PS1/467 protein.
 XX
 DE PS1/467 protein.
 XX
 KW Presentin peptide; PS1/429; immunogen; immune response; PS1 gene;
 KW Alzheimer's disease; mitochondrial pathology; neurodegeneration;
 KW apoptosis; PS1/467.
 XX
 OS Homo sapiens.
 XX
 PN WO9746678-A1.
 XX
 PD 11-DEC-1997.
 XX
 PF 03-JUN-1997; 97WO-US09272.
 XX
 PR 18-JUL-1996; 96US-0683315.
 PR

PR 06-JUN-1996; 96US-0659296.
 XX
 PA (FARB) BAYER CORP.
 XX
 PI Chisholm JC, Davis JN, Drache B;
 XX
 DR WPI: 1998-042186/04.
 DR N-PSDB; AAV17358.
 XX
 PT DNA encoding presenilin peptide PS1/429 and its analogues - useful
 PT for diagnosis and treatment of Alzheimer's disease
 XX
 PS Claim 7; Fig 2; 77pp; English.
 XX
 CC This sequence is the PS1/467 presenilin peptide. This sequence is
 CC specifically stated as not being in the nucleic acid of the invention,
 CC which encodes the PS1/429 presenilin peptide PS1/429 (II). Cells
 CC transformed with the DNA are used to produce recombinant (II) and
 CC analogues, useful e.g. as immunogens for generating an immune response
 CC against PS1/429. (II) is a new product of the PS1 gene, mutations in
 CC which cause Alzheimer's disease (AD). The nucleic acids are generally
 CC useful as probes for detection and quantification of PS1/429,
 CC particularly for diagnosis of AD, especially the target sequences that
 CC hybridise with probes are isolated for sequencing. Antibodies (Ab) can
 CC also be diagnosed at the protein level using Ab as immunoassay reagents.
 CC Ab can also be used to identify epitopes and for affinity purification of
 CC peptides. Antisense nucleic acid may also be used to regulate expression
 CC of the PS1/429 gene, and both nucleic acids and peptides are useful as
 CC size markers in electrophoresis, chromatography etc. The transgenic
 CC animals are used as models for AD, e.g. for testing drugs. Regulators of
 CC the PS1/429 gene or polypeptide can be used to treat e.g. AD or diseases
 CC involving mitochondrial pathology, apoptosis and neurodegeneration.
 CC Typical regulators are antisense sequences, ribozymes, aptamers,
 CC synthetic or natural compounds. (II) may also be used to target other
 CC coding sequences to particular cellular locations.
 XX
 SQ Sequence 467 AA;
 XX
 Query Match 100.0%; Score 15; DB 19; Length 467;
 Best Local Similarity 100.0%; Pred. No. 3.6e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SHIGPHRSTPESRAA 15
 |||||
 Db 346 SHIGPHRSTPESRAA 360

RESULT 50
 AAW41431
 ID AAW41431 standard; Protein; 467 AA.
 XX
 AC AAW41431;
 XX
 DT 04-JUN-1998 (first entry)
 XX
 DE Mouse PS1/467 protein.
 XX
 DE Presentin peptide; PS1/429; immunogen; immune response; PS1 gene;
 KW Alzheimer's disease; mitochondrial pathology; neurodegeneration;
 KW apoptosis; PS1/467.
 XX
 OS Mus sp.
 XX
 PN WO9746678-A1.
 XX
 PD 11-DEC-1997.
 XX
 PF 03-JUN-1997; 97WO-US09272.
 XX
 PR 18-JUL-1996; 96US-0683315.
 PR 06-JUN-1996; 96US-0659296.
 XX
 PA (FARB) BAYER CORP.

XX Chisholm JC, Davis JN, Drache B;
 XX
 XX WPI: 1998-042186/04.
 XX
 XX
 XX DNA encoding presenilin peptide PSI/429 and its analogues - useful
 XX for diagnosis and treatment of Alzheimer's disease
 XX
 XX Claim 7; Fig 3; 77pp; English.
 XX
 XX This sequence is the PSI/467 presenilin peptide. This sequence is
 XX specifically stated as not being in the nucleic acid of the invention,
 XX which encodes the PSI/429 presenilin peptide PSI/429 (II). Cells
 XX transformed with the DNA are used to produce recombinant (II) and
 XX analogues, useful e.g. as immunogens for generating an immune response
 XX against PSI/429. (II) is a new product of the PSI gene, mutations in
 XX which cause Alzheimer's disease (AD). The nucleic acids are generally
 XX useful as probes for detection and quantification of PSI/429,
 XX particularly for diagnosis of AD, especially the target sequences that
 XX hybridise with probes are isolated for sequencing. Antibodies (Ab) can
 XX also be diagnosed at the protein level using Ab as immunoassay reagents.
 XX Ab can also be used to identify epitopes and for affinity purification of
 XX peptides. Antisense nucleic acid may also be used to regulate expression
 XX of the PSI/429 gene, and both nucleic acids and peptides are useful as
 XX size markers in electrophoresis, chromatography etc. The transgenic
 XX animals are used as models for AD, e.g. for testing drugs. Regulators of
 XX the PSI/429 gene or polypeptide can be used to treat e.g. AD or diseases
 XX involving mitochondrial pathology, apoptosis and neurodegeneration.
 XX Typical regulators are antisense sequences, ribozymes, aptamers,
 XX synthetic or natural compounds. (II) may also be used to target other
 XX coding sequences to particular cellular locations.
 XX
 XX Sequence 467 AA:

Query Match 100.0%; Score 15; DB 19; Length 467;
 Best Local Similarity 100.0%; Pred. NO. 3.6e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SHLGPHRSTPESRAA 15
 |||||
 Db 346 SHLGPHRSTPESRAA 360

RESULT 51
 AAY24419
 ID AAY24419 standard; Protein; 467 AA.
 XX
 XX AAY24419;
 XX
 XX
 XX 23-SEP-1999 (first entry)
 XX
 XX Human presenilin-1.
 XX
 XX Presenilin-1; mutation; gene mutant animal; Alzheimer's disease.
 XX
 XX Homo sapiens.
 XX
 XX WO9334670-A1.
 XX
 XX 15-JUL-1999.
 XX
 XX 07-JAN-1999; 99WO-JP00015.
 XX
 XX 08-JAN-1998; 98JP-0002191.
 XX
 XX (DAUC) DAICHI PHARM CO LTD.
 XX
 XX Takeda J, Takeda M;
 XX
 XX WPI: 1999-430307/36.
 XX N-PSDB; AAX90184.
 XX
 XX Mutant presenilin-1 gene-introduced animals, useful as model animals

PT for study of Alzheimer's diseases in human and screening substances
 XX for prevention and/or treatment of the diseases
 XX
 XX Disclosure; Page 51-52; 64pp; Japanese.
 XX

CC The present invention describes a gene mutant animal having a non-human
 CC mutant presenilin gene. The mutant presenilin gene causes amino acid
 CC substitutions at 1 or more positions of 79, 82, 96, 115, 120, 135, 139,
 CC 143, 146, 163, 209, 213, 231, 235, 246, 260, 263, 264, 267, 269,
 CC 280, 285, 286, 290, 318, 384, 392, 410, 426 and 436, with corresponding
 CC N-terminals being e.g. A79V, V82L, A426 and P436S, particularly by
 CC replacing isoleucine of position 213 by another amino-acid especially
 CC threonine. The gene mutant animals e.g. mice can be used as model
 CC animals for the study of human Alzheimer's diseases and to screen and
 CC evaluate substances as candidates for prevention and/or therapy of
 CC Alzheimer's diseases in patients. They can over-produce amyloid
 CC beta protein by the presenilin-1 gene to cause nerve cell death or
 CC peeling off in the hippocampus earlier. Such animals are being
 CC pathologically close to human patients with Alzheimer's diseases. The
 CC present sequence represents human presenilin-1, as given in the present
 CC invention.

Sequence 467 AA:
 Query Match 100.0%; Score 15; DB 20; Length 467;
 Best Local Similarity 100.0%; Pred. NO. 3.6e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SHLGPHRSTPESRAA 15
 |||||
 Db 346 SHLGPHRSTPESRAA 360

RESULT 52
 AAY24420
 ID AAY24420 standard; Protein; 467 AA.
 XX
 XX AAY24420;
 XX
 XX
 XX 23-SEP-1999 (first entry)
 XX
 XX Mouse presenilin-1.
 XX
 XX Presenilin-1; mutation; gene mutant animal; Alzheimer's disease.
 XX
 XX Mus sp.
 XX
 XX WO9334670-A1.
 XX
 XX 15-JUL-1999.
 XX
 XX 07-JAN-1999; 99WO-JP00015.
 XX
 XX 08-JAN-1998; 98JP-0002191.
 XX
 XX (DAUC) DAICHI PHARM CO LTD.
 XX
 XX Takeda J, Takeda M;
 XX
 XX WPI: 1999-430307/36.
 XX N-PSDB; AAX90185.
 XX
 XX Mutant presenilin-1 gene-introduced animals, useful as model animals
 XX for prevention and/or treatment of the diseases
 XX
 XX Disclosure; Page 54-56; 64pp; Japanese.
 XX
 XX The present invention describes a gene mutant animal having a non-human
 XX mutant presenilin gene. The mutant presenilin gene causes amino acid
 XX substitutions at 1 or more positions of 79, 82, 96, 115, 120, 135, 139,
 XX 143, 146, 163, 209, 213, 231, 235, 246, 250, 260, 263, 264, 267, 269,
 XX 280, 285, 286, 290, 318, 384, 392, 410, 426 and 436, with corresponding

CC N-terminals being e.g. A79V, V82L, A426 and P436S, particularly by
 CC replacing isoleucine of position 213 by another amino-acid especially
 CC threonine. The gene mutant animals e.g. mice can be used as model
 CC animals for the study of human Alzheimer's diseases and to screen and
 CC evaluate substances as candidates for prevention and/or therapy of
 CC Alzheimer's diseases in patients. They can over-produce amyloid
 CC beta protein by the presenilin-1 gene to cause nerve cell death or
 CC peeling off in the hippocampus earlier. Such animals are being
 CC pathologically close to human patients with Alzheimer's diseases. The
 CC present sequence represents mouse presenilin-1, as given in the present
 CC invention.

XX Sequence 467 AA;

Query Match 100.0%; Score 15; DB 20; Length 467;

Best Local Similarity 100.0%; Pred. No. 3.6e-08; Mismatches 0; Gaps 0;

Matches 15; Conservative 0; Indels 0; Gaps 0;

OY 1 SHLGPHRSTPESRAA 15

DB 346 SHLGPHRSTPESRAA 360

RESULT 53

AAZ23897

ID AAZ23897 standard; protein; 467 AA.

XX AAZ23897;

DT 27-SEP-1999 (first entry)

DE Amino acid sequence of the wild type human presenilin 1 (PS1) protein.

XX Human; Presenilin 1; PS1; presenilin-binding protein; interacting domain;

KW presenilin allele; Alzheimer's disease; senile dementia;

KW psychiatric disease; schizophrenia; depression; neurological disease;

KW stroke; cerebral haemorrhage.

XX Homo sapiens.

XX W09935501-A1.

XX 15-JUL-1999.

XX 08-JAN-1999; 99WO-CA00018.

XX 09-JAN-1998; 98US-0070948.

XX (UTOR) UNIV TORONTO GOVERNING COUNCIL.

XX Fraser PE) St George-Hyslop PH;

XX WPI: 1999-419410/35.

XX Identifying substances that alter presenilin interactions, useful
 PT for screening individuals for presenilin alleles associated with
 PT Alzheimer's disease - useful for diagnosis of Alzheimer's disease

XX Disclosure: Page 34-36; 40pp; English.

XX The present sequence represents wild type human presenilin 1 (PS1)
 CC protein. The specification describes a method for identifying substances
 CC that alter the interaction of a presenilin with a presenilin-binding
 CC protein. The method comprises contacting the interacting domain of a
 CC presenilin protein to a presenilin-binding protein in the presence of a
 CC test substance, and measuring the interaction of the presenilin and
 CC the presenilin-binding protein. The method can be used to screen
 CC individuals for presenilin alleles associated with Alzheimer's disease
 CC and related disorders, such as senile dementia's, psychiatric diseases
 CC such as schizophrenia and depression, and neurological disease, such as
 CC stroke and cerebral haemorrhage.

XX Sequence 467 AA;

Query Match 100.0%; Score 15; DB 20; Length 467;
 Best Local Similarity 100.0%; Pred. No. 3.6e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGPHRSTPESRAA 15

DB 346 SHLGPHRSTPESRAA 360

RESULT 54

AAB07971

ID AAB07971 standard; protein; 467 AA.

XX AAB07971;

DT 14-NOV-2000 (first entry)

DE Amino acid sequence of human presenilin 1 polypeptide.

XX Human; presenilin; PS1; PS2; nerve cell growth; nerve damage; hNRPAP;
 KW neural plakophilin related armadillo protein; Alzheimer's disease;
 KW Parkinson's disease; amyotrophic lateral sclerosis; ALS; stroke;
 KW multiple sclerosis; ischemia; stroke; neural paropathy; sciatic crush;
 KW motor neuron disease; peripheral neuropathy; neuropathy; diabetes;
 KW spinal cord injury; facial nerve crush.

XX Homo sapiens.

XX W020047615-A2.

XX 17-AUG-2000.

XX 11-FEB-2000; 2000WO-CA00126.

XX 12-FEB-1999; 99US-0119835.

XX (UTOR) UNIV TORONTO GOVERNING COUNCIL.

XX St George-Hyslop PH, Fraser PE;

XX WPI: 2000-524531/47.

XX Stimulation of nerve cell growth using human Neural Plakophilin Related
 PT Armadillo Protein (hNRPAP) polypeptide, useful for the treatment of
 PT diseases such as Alzheimer's, Parkinson's, and stroke -

XX Disclosure: Page 19-20; 33pp; English.

XX The present sequence represents a human presenilin 1 (PS1) polypeptide.
 CC Human Neural Plakophilin Related Armadillo Protein (hNRPAP) polypeptide
 CC is known to interact with PS1 and PS2. The specification describes a
 CC method for stimulating the growth of nerve cells, comprising contacting
 CC them with hNRPAP. The hNRPAP polypeptide and polynucleotide are useful
 CC for treating nerve damage caused by a variety of diseases or physical
 CC traumas, including Alzheimer's disease, Parkinson's disease, amyotrophic
 CC lateral sclerosis (ALS), multiple sclerosis, stroke, ischemia associated
 CC with stroke, neural paropathy, motor neuron diseases, sciatic crush,
 CC peripheral neuropathy, neuropathy associated with diabetes, spinal cord
 CC injuries and facial nerve crush.

XX Sequence 467 AA;

Query Match 100.0%; Score 15; DB 21; Length 467;

Best Local Similarity 100.0%; Pred. No. 3.6e-08; Mismatches 0; Gaps 0;

Matches 15; Conservative 0; Indels 0; Gaps 0;

OY 1 SHLGPHRSTPESRAA 15

DB 346 SHLGPHRSTPESRAA 360

RESULT 55

AAE10798
ID AAE10798 standard; Protein; 467 AA.
XX
AC AAE10798;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human presenilin-1 (PS-1) protein.
XX
KW Human: catenin p120; presenilin-1; PS-1; neuroprotective; gene therapy;
KW neurodegenerative disease; Alzheimer's disease; nootropic; prophylaxis;
KW neuronal disorder; cognitive disorder.
XX
OS Homo sapiens.
XX
PN WO200167097-A2.
XX
PD 13-SEP-2001.
XX
PF 09-MAR-2001; 2001WO-GB01059.
XX
PR 10-MAR-2000; 2000GB-0005895.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Hale RS, Rowley A, Blackstock W;
XX
DR WPI; 2001-589954/66.
DR N-PSDB; AAD18120.
XX
PT Identifying presenilin or catenin p120 activity modulator useful for
PT modulating presenilin-catenin p120 interaction and thus for treating
PT cognitive disorder e.g., Alzheimer's disease comprises enhancing
PT cognitive function -
XX
PS Example 1; Page 41-42; 48pp; English.
XX
CC The invention relates to a method for identifying modulators of
CC presenilin and catenin p120. Modulators of catenin p120 and presenilin
CC are useful for the treatment and prophylaxis of disorders that is
CC responsive to modulation of presenilin/catenin p120 activity. In
CC particular, neuronal disorders such as cognitive disorders and
CC neurodegenerative diseases such as Alzheimer's disease. Catenin p120 DNAs
CC are useful for identifying mutations in catenin p120 genes.
CC Identification of such mutations assist in the diagnosis of or
CC susceptibility to Alzheimer's or other conditions associated with
CC presenilin and in assessing the physiology of such disorders. Catenin
CC p120 DNAs are also used in hybridisation studies to monitor expression of
CC p120 genes and in particular for up or down regulation of catenin p120
CC expression. The present sequence is human presenilin-1 (PS-1) protein.
XX
SQ Sequence 467 AA;
XX
Query Match 100.0%; Score 15; DB 22; Length 467;
Best Local Similarity 100.0%; Pred. No. 3.6e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 SHLGHSTPESRAA 15
Db 346 SHLGHSTPESRAA 360
XXXXXXXXXXXXXXXXXXXX
Db 346 SHLGHSTPESRAA 360
XXXXXXXXXXXXXXXXXXXX
XX
RESULT 56
AAE05466
ID AAE05466 standard; Protein; 467 AA.
XX
AC AAE05466;
XX
DT 29-OCT-2001 (first entry)
XX
DE Amino acid sequence of human presenilin 1.
XX
KW KIAA0253; presenilin; Alzheimer's disease.

XX
OS Homo sapiens.
XX
PN WO200167109-A1.
XX
PD 13-SEP-2001.
XX
PF 09-MAR-2001; 2001WO-GB01057.
XX
PR 10-MAR-2000; 2000GB-0005894.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Hale RS, Rowley A, Blackstock W;
XX
DR WPI; 2001-522960/57.
DR N-PSDB; AAH74993.
XX
PT Identifying a modulator of presenilin function by determining the
PT ability of presenilin to bind to a KIAA0253 polypeptide in the presence
PT and absence of a test compound, useful in the treatment or prophylaxis
PT of Alzheimer's disease -
XX
PS Disclosure; Page 40-41; 48pp; English.
XX
CC The present sequence represents human presenilin 1. KIAA0253 binds to
CC presenilin. The specification describes a method of identifying a
CC modulator of presenilin function or KIAA0253 function. The method
CC comprises determining presenilin activity or KIAA0253 activity
CC in the presence and absence of a test compound, where presenilin
CC activity is determined by its ability to bind to KIAA0253. A modulator
CC of presenilin or KIAA0253 polypeptide is useful in the manufacture
CC of a medicament for the treatment or prophylaxis of Alzheimer's disease.
CC The KIAA0253 polynucleotide and KIAA0253 polypeptide are useful in the
CC treatment, prophylaxis or diagnosis of Alzheimer's disease.
XX
SQ Sequence 467 AA;
XX
Query Match 100.0%; Score 15; DB 22; Length 467;
Best Local Similarity 100.0%; Pred. No. 3.6e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 SHLGHSTPESRAA 15
Db 346 SHLGHSTPESRAA 360
XXXXXXXXXXXXXXXXXXXX
Db 346 SHLGHSTPESRAA 360
XXXXXXXXXXXXXXXXXXXX
XX
RESULT 57
AAE05466
ID AAE05466 standard; Protein; 467 AA.
XX
AC AAE05466;
XX
DT 24-SEP-2001 (first entry)
XX
DE Human presenilin (PS1) protein.
XX
KW Human: Par-4; presenilin; PS1; neuroprotective; nuclear factor kappa B;
KW NF-kappa B; neuronal degeneration; spinal muscular atrophy; paralysis;
KW peripheral neuropathy; motor neuron disorder; neurodegenerative disorder;
KW Parkinson's disease; Meniere's disease; multiple sclerosis; Bell's palsy;
KW Huntington's chorea; Down's syndrome; amyotrophic lateral sclerosis; ALS;
KW nerve deafness; Alzheimer's disease; epilepsy.
XX
OS Homo sapiens.
XX
PN WO200151671-A2.
XX
PD 19-JUL-2001.
XX
PF 08-JAN-2001; 2001WO-US00526.
XX
PR 10-JAN-2000; 2000US-0175200.

PR 04-JAN-2001; 2001US-0754949.
XX
XX (SCIO-) SCIOS INC.
XX
XX McCarthy J, Cordell B;
XX
XX MPI; 2001-451872/48.
XX
XX N-PSDB; AAE05466.
XX
XX
XX Identifying inhibitors of neuronal degeneration useful for treating
XX e.g. Alzheimer's disease, by determining the ability of a compound to
XX induce nuclear factor kappa B activation, with the involvement of
XX presenilin or Par-4
XX
XX Claim 4; Page 60-61; 66pp; English.
XX
XX The invention relates to human Par-4 protein, presenilin protein (PS1
XX and PS2) and their corresponding DNA molecules. The invention also
XX relates to a method for identifying inhibitors of neuronal degeneration,
XX comprising cotransfecting eukaryotic host cells expressing presenilin
XX (PS), with a Par-4 DNA, and an NF-kappa B dependent reporter construct,
XX exposing the cotransfected cells to a candidate molecule and monitoring
XX the ability of the candidate molecule to induce NF-kappa B activation.
XX Presenilin proteins participate in nuclear factor kappa B (NF-kappa B)
XX signaling and activation. The inhibitors of neuronal degeneration
XX are useful for treating neurodegenerative disorders such as Alzheimer's
XX disease, Parkinson's disease, epilepsy, multiple sclerosis, Huntington's
XX chorea, Down's syndrome, nerve deafness, Meniere's disease and also for
XX treating peripheral neuropathies, motorneuron disorders such as
XX amyotrophic lateral sclerosis (ALS), Bell's palsy and various conditions
XX involving spinal muscular atrophy and paralysis. The present sequence
XX is human presenilin (PS1) protein.
XX
XX Sequence 467 AA;
XX
XX
XX Query Match 100.0%; Score 15; DB 22; Length 467;
XX Best Local Similarity 100.0%; Pred. No. 3.6e-08;
XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 SHLGPHRSTPESRAA 15
XX 346 SHLGPHRSTPESRAA 360
XX
XX
XX RESULT 58
XX AAE05563
XX ID AAE05563 standard; Protein; 467 AA.
XX
XX AAE05563;
XX
XX 24-SEP-2001 (first entry)
XX
XX Human presenilin PS1-FAD mutant M146V.
XX
XX Human; Par-4; presenilin; PS1; neuroprotective; nuclear factor kappa B;
XX NF-kappa B; neuronal degeneration; spinal muscular atrophy; paralysis;
XX peripheral neuropathy; motorneuron disorder; neurodegenerative disorder;
XX Parkinson's disease; Meniere's disease; multiple sclerosis; Bell's palsy;
XX Huntington's chorea; Down's syndrome; amyotrophic lateral sclerosis; ALS;
XX nerve deafness; Alzheimer's disease; epilepsy; mutant; mutein.
XX
XX Homo sapiens.
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Misc-difference 146 /note="Wild type Met substituted with Val"
XX
XX MO200151671-A2.
XX
XX 19-JUL-2001.
XX
XX 08-JAN-2001; 2001WO-US00526.
XX
XX PF

XX
XX 10-JAN-2000; 2000US-0175200.
XX
XX 04-JAN-2001; 2001US-0754949.
XX
XX (SCIO-) SCIOS INC.
XX
XX McCarthy J, Cordell B;
XX
XX MPI; 2001-451872/48.
XX
XX
XX Identifying inhibitors of neuronal degeneration useful for treating
XX e.g. Alzheimer's disease, by determining the ability of a compound to
XX induce nuclear factor kappa B activation, with the involvement of
XX presenilin or Par-4
XX
XX Example 2; Page -: 66pp; English.
XX
XX The invention relates to human Par-4 protein, presenilin protein (PS1
XX and PS2) and their corresponding DNA molecules. The invention also
XX relates to a method for identifying inhibitors of neuronal degeneration,
XX comprising cotransfecting eukaryotic host cells expressing presenilin
XX (PS), with a Par-4 DNA, and an NF-kappa B dependent reporter construct,
XX exposing the cotransfected cells to a candidate molecule and monitoring
XX the ability of the candidate molecule to induce NF-kappa B activation.
XX Presenilin proteins participate in nuclear factor kappa B (NF-kappa B)
XX signaling and activation. The inhibitors of neuronal degeneration
XX are useful for treating neurodegenerative disorders such as Alzheimer's
XX disease, Parkinson's disease, epilepsy, multiple sclerosis, Huntington's
XX chorea, Down's syndrome, nerve deafness, Meniere's disease and also for
XX treating peripheral neuropathies, motorneuron disorders such as
XX amyotrophic lateral sclerosis (ALS), Bell's palsy and various conditions
XX involving spinal muscular atrophy and paralysis. The present sequence
XX is human presenilin PS1-FAD (familial Alzheimer's disease) mutant.
XX Note: This sequence is not shown in the specification but is derived from
XX human presenilin (PS1) protein [SEQ ID NO: 4] shown in page 60-61 of the
XX specification (AAE05466).
XX
XX Sequence 467 AA;
XX
XX
XX Query Match 100.0%; Score 15; DB 22; Length 467;
XX Best Local Similarity 100.0%; Pred. No. 3.6e-08;
XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 SHLGPHRSTPESRAA 15
XX 346 SHLGPHRSTPESRAA 360
XX
XX
XX RESULT 59
XX AAE05564
XX ID AAE05564 standard; Protein; 467 AA.
XX
XX AAE05564;
XX
XX 24-SEP-2001 (first entry)
XX
XX Human presenilin PS1-FAD mutant E280G.
XX
XX Human; Par-4; presenilin; PS1; neuroprotective; nuclear factor kappa B;
XX NF-kappa B; neuronal degeneration; spinal muscular atrophy; paralysis;
XX peripheral neuropathy; motorneuron disorder; neurodegenerative disorder;
XX Parkinson's disease; Meniere's disease; multiple sclerosis; Bell's palsy;
XX Huntington's chorea; Down's syndrome; amyotrophic lateral sclerosis; ALS;
XX nerve deafness; Alzheimer's disease; epilepsy; mutant; mutein.
XX
XX Homo sapiens.
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Misc-difference 280 /note="Wild type Glu substituted with Gly"
XX
XX MO200151671-A2.
XX
XX FT
XX PN

XX 19-JUN-2001.
PD 08-JAN-2001; 2001MO-US00526.
XX 10-JAN-2000; 2000US-0175200.
PR 04-JAN-2000; 2001US-0754949.
XX (SCIO-) SCIOS INC.
PA McCarthy J, Cordell B;
XX MPI; 2001-451872/48.
DR
XX
PT Identifying inhibitors of neuronal degeneration useful for treating
PT e.g. Alzheimer's disease, by determining the ability of a compound to
PT induce nuclear factor kappa B activation, with the involvement of
PT presenilin or Par-4
XX
XX Example 2; Page -: 66pp; English.
PS
XX The invention relates to human Par-4 protein, presenilin protein (PS1
CC and PS2) and their corresponding DNA molecules. The invention also
CC relates to a method for identifying inhibitors of neuronal degeneration,
CC comprising cotransfecting eukaryotic host cells expressing presenilin
CC (PS), with a Par-4 DNA, and an NF-kappa B dependent reporter construct,
CC exposing the cotransfected cells to a candidate molecule and monitoring
CC the ability of the candidate molecule to induce NF-kappa B activation.
CC Presenilin protein participates in nuclear factor kappa B (NF-kappa B)
CC signaling and activation. The inhibitors of neuronal degeneration
CC are useful for treating neurodegenerative disorders such as Alzheimer's
CC disease, Parkinson's disease, epilepsy, multiple sclerosis, Huntington's
CC chorea, Down's syndrome, nerve deafness, Meniere's disease and also for
CC treating peripheral neuropathies, motor neuron disorders such as
CC amyotrophic lateral sclerosis (ALS), Bell's palsy and various conditions
CC involving spinal muscular atrophy and paralysis. The present sequence
CC is human presenilin PS1-PAD (familial Alzheimer's disease) mutant.
CC Note: This sequence is not shown in the specification but is derived from
CC human presenilin (PS1) protein [SEQ ID NO: 4] shown in page 60-61 of the
CC specification (AA05466).
XX
SQ Sequence 467 AA:
Query Match 100.0%; Score 15; DB 22; Length 467;
Best Local Similarity 100.0%; Pred. No. 3.6e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SHLGPHRSTPESRAA 15
DB 346 SHLGPHRSTPESRAA 360
RESULT 60
AA018049
ID AA018049 standard; Protein; 467 AA.
XX
AC AA018049;
XX
DT 02-SEP-2002 (first entry)
XX
DE Presenilin protein.
XX
KM Sel-12; presenilin; neuronal disorder; familial Alzheimer's disease;
XX amyloid precursor protein; APP.
XX
OS Unidentified.
XX
PN US6376239-B1.
XX
PD 23-APR-2002.
PF 04-APR-1997; 97US-0832867.
XX

PR 04-APR-1997; 97US-0832867.
XX (ELEG-) ELEGENE GMBH.
PA
XX
XX Baumeister R;
XX
DR WPI; 2002-478281/51.
DR N-PSDB; AAL47323.
XX
PT Isolated DNA molecule comprising promoter of the sel-12 gene from
PT Caenorhabditis elegans operably linked to heterologous gene, directs
PT expression in neural cells and is useful to develop drugs to treat
PT neuronal disorders
XX
XX Claim 2; Fig 4; 78pp; English.
PS
XX The present invention relates to DNA molecules comprising the promoter of
CC the sel-12 gene from Caenorhabditis elegans operably linked to a
CC heterologous DNA sequence encoding a protein of interest. The sequence
CC can be used to develop drugs for the treatment, prevention or delay of a
CC neuronal disorder. In particular, the neuronal disorder may be familial
CC Alzheimer's disease. The present sequence is a presenilin protein
CC described in the exemplification of the invention.
XX
SQ Sequence 467 AA:
Query Match 100.0%; Score 15; DB 23; Length 467;
Best Local Similarity 100.0%; Pred. No. 3.6e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SHLGPHRSTPESRAA 15
DB 346 SHLGPHRSTPESRAA 360
RESULT 61
AA079416
ID AA079416 standard; protein; 467 AA.
XX
AC AA079416;
XX
DT 02-JUL-2002 (first entry)
XX
DE Human presenilin-1, PS-1.
XX
KW Human; integrin-linked kinase; ILK; presenilin-1; protein kinase B; PKB;
KW gamma secretase; apoptosis; Alzheimer's disease; PS-1; nootropic;
KW neuroprotective; cyostatic; cancer.
XX
OS Homo sapiens.
XX
PN W0200222862-A2.
XX
PD 21-MAR-2002.
XX
PF 12-SEP-2001; 2001MO-GB04094.
XX
PR 12-SEP-2000; 2000GB-0022333.
XX
PA (GLAXO) GLAXO GROUP LTD.
XX
PI Hiles ID, Ellis C;
XX
DR WPI; 2002-351896/38.
XX
PT Identifying agent that modulates interaction between integrin-linked
PT kinase and presenilin-1, useful for treating Alzheimer's disease, and
PT identifying agent that modulates protein kinase B or gamma secretase
PT activity
XX
PS Disclosure; Page 43-45; 53pp; English.
XX
CC The invention relates to identifying an agent modulating interaction

XX	30-JUN-2000; 2000US-215345P.	
PR		
PA	(PHAA) PHARMACIA & UPJOHN CO.	
XX		
PI	Carter DB, Tomasselli AG;	
XX		
DR	WPI: 2002-140082/18.	
DR	N-PSDB; AAD27444.	
XX		
PS	Novel isolated mutant presenilin 1 and presenilin 2 polypeptides, with	
PT	useful for screening of drugs for treating pathologies associated with	
PT	aberrant amyloid precursor protein processing, such as Alzheimer's	
PT	disease	
XX		
PS	Claim 15; Page 68-70; 80pp; English.	
XX		
CC	The invention relates to mutant presenilin 1 (PS1) and presenilin 2	
CC	(PS2) polypeptides. Presenilin are involved in the processing of amyloid	
CC	precursor protein (APP) from which major amyloidogenic peptides are	
CC	cleaved. Mutant presenilins are useful for identifying agents that	
CC	modulate amyloid beta-peptide (Abeta) derived peptide production. Mutant	
CC	presenilin 1 is also useful as a target for screening drugs useful in the	
CC	treatment of pathologies associated with aberrant amyloid precursor	
CC	protein processing, such as Alzheimer's disease, Parkinson's disease,	
CC	multiple sclerosis, Huntington's disease, amyotrophic lateral sclerosis,	
CC	head injury disease, Picks disease, frontal lobe dementia, cerebellar	
CC	degeneration, stroke, ischaemic injury and schizophrenia. A transgenic	
CC	non-human animal is useful for analysing the interaction between APP and	
CC	mutant presenilin-processing protease in vivo, and for screening anti-	
CC	Alzheimer's disease drugs in vivo. A transgenic non-human	
CC	animal is useful for analysing the interaction between APP and mutant	
CC	presenilin-processing protease in vivo, and for screening anti-	
CC	Alzheimer's disease drugs in vivo. The present sequence is human	
CC	mutant PS1 protein.	
XX		
SO	Sequence 467 AA;	
XX		
OY	Query Match 100.0%; Score 15; DB 23; Length 467;	
DB	Best Local Similarity 100.0%; Prid. NO. 3.6e-08;	
	Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
	1 SHLGPHRSTPESRAA 15	
	346 SHLGPHRSTPESRAA 360	
RESULT 64		
AAEI17047		
ID	AAEI17047 standard; Protein: 467 AA.	
XX		
AC	AAEI17047;	
XX		
DT	18-APR-2002 (first entry)	
XX		
XX	Human mutant presenilin 1 (PS1) protein #3.	
KW	Human; presenilin 1; PS1; amyloid precursor protein; APP; drug screening;	
KW	Alzheimer's disease; Parkinson's disease; multiple sclerosis; stroke;	
KW	Huntington's disease; amyotrophic lateral sclerosis; Picks disease;	
KW	head injury disease; frontal lobe dementia; cerebellar degeneration;	
KW	Ischaemic injury; schizophrenia; mutant; muteln.	
XX		
OS	Homo sapiens.	
OS	Synthetic.	
XX		
FT	Key Location/Qualifiers	
FT	Misc-difference 207..210	
FT	/note= "Encoded by GTGCGTGC"	
FT	Misc-difference 258	
FT	/label= Unknown	
FT	/note= "Wild type Leu substituted with Xaa; Encoded	
FT	by NNN"	

```

FT Misc-difference 259 /Label= Unknown
FT /note= "Wild type Val substituted with Xaa; Encoded by NNN"
FT MISC-difference 386 /Label= Unknown
FT /note= "Wild type Phe substituted with Xaa; Encoded by NNN"
FT Misc-difference 387 /Label= Unknown
FT /note= "Wild type Ile substituted with Xaa; Encoded by NNN"
XX W0200202601-A2.
XX 10-JAN-2002.
XX 29-JUN-2001; 2001WO-US16508.
XX 30-JUN-2000; 2000US-215345P.
XX (PHAA ) PHARMACIA & UPJOHN CO.
XX Carter DB, Tomasselli AG;
XX WPI: 2002-140082/18.
XX N-PSDB; AAD27446.
XX Novel isolated mutant presenilin 1 and presenilin 2 polypeptides, useful for screening of drugs for treating pathologies associated with aberrant amyloid precursor protein processing, such as Alzheimer's disease -
XX Claim 67; Page 74-75; 80pp; English.
XX The invention relates to mutant presenilin 1 (PS1) and presenilin 2 (PS2) polypeptides. Presenilin are involved in the processing of amyloid precursor protein (APP) from which major amyloidogenic peptides are cleaved. Mutant presenilins are useful for identifying agents that modulate amyloid beta-peptide (Abeta) derived peptide production. Mutant presenilin is also useful as a target for screening drugs useful in the treatment of pathologies associated with aberrant amyloid precursor protein processing, such as Alzheimer's disease, Parkinson's disease, multiple sclerosis, Huntington's disease, amyotrophic lateral sclerosis, head injury disease, Pick's disease, frontal lobe dementia, cerebellar degeneration, stroke, ischemic injury and schizophrenia. A transgenic non-human animal is useful for analysing the interaction between APP and mutant presenilin-processing protease in vivo, and for screening anti-Alzheimer's disease drugs in vivo. A transgenic non-human animal is useful for analysing the interaction between APP and mutant presenilin-processing protease in vivo, and for screening anti-Alzheimer's disease drugs in vivo. The present sequence is human mutant PS1 cDNA.
SO Sequence 467 AA:
OY 1 SHLGPHRSTPESRAA 15 |||||
DB 346 SHLGPHRSTPESRAA 360 |||||
Query Match 100.0%; Score 15; DB 23; Length 467;
Best Local Similarity 100.0%; Pred. No. 3.6e+08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
RESULT 65
AAEI7051
ID AAEI7051 standard; Protein: 467 AA.
XX AAEI7051;
XX 18-APR-2002 (first entry)

```

DE Human mutant presenilin 1 (PS1) wild type protein.
 XX
 KW Human; presenilin 1; PS1; amyloid precursor protein; APP; drug screening;
 OS Alzheimer's disease; Parkinson's disease; multiple sclerosis; stroke;
 XX Huntington's disease; amyotrophic lateral sclerosis; Pick's disease;
 KW head injury disease; frontal lobe dementia; cerebellar degeneration;
 XX ischaemic injury; schizophrenia.
 XX Homo sapiens.
 OS
 XX MO200202601-A2.
 PN
 XX 10-JAN-2002.
 PD
 XX 29-JUN-2001; 2001WO-US16508.
 PF
 XX 30-JUN-2000; 2000US-215345P.
 PR
 XX (PHAA) PHARMACIA & UPJOHN CO.
 PA
 XX Carter DB, Tomasselli AG;
 PI
 XX MPI; 2002-140082/18.
 DR
 XX
 PT Novel isolated mutant presenilin 1 and presenilin 2 polypeptides,
 PT useful for screening of drugs for treating pathologies associated with
 PT aberrant amyloid precursor protein processing, such as Alzheimer's
 PT disease -
 PS
 XX Disclosure; Fig 1; 80pp; English.
 XX
 CC The invention relates to mutant presenilin 1 (PS1) and presenilin 2
 CC (PS2) polypeptides. Presenilin are involved in the processing of amyloid
 CC precursor protein (APP) from which major amyloidogenic peptides are
 CC cleaved. Mutant presenilins are useful for identifying agents that
 CC modulate amyloid beta-peptide (Abeta) derived peptide production. Mutant
 CC presenilin is also useful as a target for screening drugs useful in the
 CC treatment of pathologies associated with aberrant amyloid precursor
 CC protein processing, such as Alzheimer's disease, Parkinson's disease,
 CC multiple sclerosis, Huntington's disease, amyotrophic lateral sclerosis,
 CC head injury disease, Pick's disease, frontal lobe dementia, cerebellar
 CC degeneration, stroke, ischemic injury and schizophrenia. A transgenic
 CC non-human animal is useful for analysing the interaction between APP and
 CC mutant presenilin-processing protease in vivo, and for screening anti-
 CC Alzheimer's disease drugs in vivo. A transgenic non-human
 CC animal is useful for analysing the interaction between APP and mutant
 CC presenilin-processing protease in vivo, and for screening anti-
 CC Alzheimer's disease drugs in vivo. The present sequence is human
 CC PSI wild type protein.
 CC
 SQ Sequence 467 AA;
 XX
 XX
 Query Match 100.0%; Score 15; DB 23; Length 467;
 Best Local Similarity 100.0%; Pred. No. 3.6e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 SHLGPHRSTPESRAA 15
 DB 346 SHLGPHRSTPESRAA 360
 ID AAM05785 standard; peptide: 14 AA.
 AC AAM05785;
 XX
 XX 28-JUL-1997 (first entry)
 DT
 XX
 DE Presentin-1-1 residues 346-359.
 XX
 KW Presentin-1; human; hps1-1; hps1-2; PS-2; integral membrane protein; AD;
 KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;

KW depression; antibody; gene expression modulator; therapy; mutein.
 XX
 OS Homo sapiens.
 XX
 XX MO9634099-A2.
 PN
 XX 31-OCT-1996.
 PD
 XX 29-APR-1996; 96WO-CA00263.
 PF
 XX 31-JUL-1995; 95US-0509359.
 PR 28-APR-1995; 95US-0431048.
 PR 28-JUN-1995; 95US-0496841.
 XX
 XX (HSCR-) HSC RES & DEV LP.
 PA (UTOR) UNIV TORONTO GOVERNING COUNCIL.
 XX
 XX Fraser PE, Rommens JM, St George-Hyslop PH;
 PI
 XX MPI; 1996-497631/49.
 DR
 XX
 XX
 PT New presenilin genes - useful for diagnosis, therapy and drug
 PT screening of familial Alzheimer's disease, cerebral disorders, etc.
 PT
 XX Claim 71; Page -; 178pp; English.
 PS
 XX AAM05768-M05788 represent antigenic fragments of the human
 XX presenilin-1-1 protein (see AAM05733 for wild type sequence). AAM05734
 XX represents a different wild type form of presenilin-1 that results from
 XX alternate splicing of the genomic DNA sequence. The presenilins are a
 XX family of highly conserved integral membrane proteins with a common
 XX structural motif, common alternate splicing patterns, and common
 XX mutational hot spot regions. Mutations in PS genes are implicated in
 XX familial Alzheimer's disease (AD) and possibly other diseases such as
 XX cerebral haemorrhage, schizophrenia, depression etc., so detection of
 XX mutations in the DNA encoding the wild type sequences can be used for
 XX diagnosis of these diseases. The wild type proteins, or vectors that
 XX express them or containing antisense sequences, antibodies selective for
 XX these mutant forms of the proteins and modulators of PS gene expression
 XX are potentially useful for treatment of AD etc. Transgenic animals are
 XX useful as models for drug screening. The antibodies can also be used e.g.
 XX for affinity purification and in immunoassays.
 XX
 SQ Sequence 14 AA;
 XX
 XX
 Query Match 93.3%; Score 14; DB 17; Length 14;
 Best Local Similarity 100.0%; Pred. No. 2e-08;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 SHLGPHRSTPESRA 14
 DB 1 SHLGPHRSTPESRA 14
 ID AAM34093 standard; peptide: 16 AA.
 AC AAM34093;
 XX
 XX 05-MAY-1998 (first entry)
 DT
 XX
 DE Peptide derived from the C-terminal of the Presentin-1 protein.
 XX
 XX Presentin-1; antibody; cleavage; cleavage inhibition; treatment;
 KW prevention; Alzheimer's disease; cleavage assay; prognosis.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 XX MO9741443-A2.
 PN
 XX 06-NOV-1997.
 PD

XX 22-APR-1997; 97WO-EP02050.
 XX
 XX 26-APR-1996; 96GB-0008657.
 XX
 XX (SMK) SMITHLINE BEECHAM PLC.
 XX
 XX Davis JB, Gray C, Karran EH, Ward RV;
 XX WPI: 1997-549892/50.
 XX
 XX Identifying compounds that inhibit cleavage of presenilin-1 - for
 XX treatment, prevention and prognosis of Alzheimer's disease
 XX
 XX Claim 9; Page 15; 17pp; English.
 XX
 XX The present peptide is derived from the C-terminal of the presenilin-1
 XX protein, amino acid residues 344-358. Rabbit polyclonal antibodies were
 XX raised against this peptide. This antibody was used, together with an
 XX antibody raised against a peptide derived from the N-terminal, to
 XX demonstrate cleavage of presenilin-1 in cells. A novel method for
 XX screening compounds that modulate, especially inhibit, cleavage of
 XX presenilin-1 comprises measuring their effect on cleavage of presenilin-1
 XX to its 18 kD and 28 kD fragments. These compounds are potentially useful
 XX for treatment and prevention of Alzheimer's disease. Assays for cleavage
 XX can be used for prognosis of Alzheimer's disease.
 XX
 XX Sequence 16 AA:
 SQ
 Query Match 86.7%; Score 13; DB 18; Length 16;
 Best Local Similarity 100.0%; Pred. No. 2.6e-07;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SHLGPHRSTPEER 13
 |||||
 DB 4 SHLGPHRSTPEER 16
 RESULT 68
 AAE12898
 ID AAE12898 standard; peptide; 16 AA.
 XX
 XX AAE12898;
 AC
 XX
 XX 15-JAN-2002 (first entry)
 DT
 XX
 XX Presenilin peptide, PSI #1.
 DE
 XX
 XX Alzheimer's disease; gamma-secretase; integral-membrane protein;
 KW beta-amyloid precursor protein; betaAPP; presenilin; PSI.
 KM
 XX
 XX Synthetic.
 OS
 XX
 XX Key Location/Qualifiers
 FH Modified-site 16
 FT /note="C-terminal amide"
 FT
 XX
 XX WO2001/5435-A2.
 PN
 XX
 XX 11-OCT-2001.
 PD
 XX
 XX 30-MAR-2001; 2001WO-US10453.
 PF
 XX
 XX 03-APR-2000; 2000US-194495P.
 PR
 XX
 XX (BRIM) BRISTOL-MYERS SQUIBB CO.
 PA
 XX
 XX Roberts SB, Hendrick JP, Vinitzky A, Lewis M, Smith DW, Pak R;
 PI WPI: 2001-648575/74.
 DR
 XX
 XX Novel gamma secretase protein, useful in the production of amyloids, is
 PT capable of cleaving beta-amyloid precursor protein to produce beta

PT amyloid peptide -
 XX
 XX Example 8; Page 26; 127pp; English.
 PS
 XX The invention relates to the field of plaque amyloid deposits that are
 CC the hallmarks of Alzheimer's disease. In particular, the invention
 CC relates to an isolated, functionally-active protein that has
 CC gamma-secretase activity. Gamma-secretase activity is necessary for
 CC amyloid production. The present invention also relates to methods for
 CC isolating integral-membrane proteins and protein complexes, including
 CC the gamma-secretase protein of the invention. The method is useful for
 CC monitoring the cleavage of beta-amyloid precursor protein (betaAPP)
 CC by gamma-secretase. The present sequence is a synthetic peptide
 CC antigen of presenilin peptide, PSI. This sequence is used in the
 CC exemplification of the invention.
 XX
 XX Sequence 16 AA:
 SQ
 Query Match 86.7%; Score 13; DB 22; Length 16;
 Best Local Similarity 100.0%; Pred. No. 2.6e-07;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SHLGPHRSTPEER 13
 |||||
 DB 4 SHLGPHRSTPEER 16
 RESULT 69
 AAE12902
 ID AAE12902 standard; peptide; 16 AA.
 XX
 XX AAE12902;
 AC
 XX
 XX 15-JAN-2002 (first entry)
 DT
 XX
 XX Presenilin peptide, PSI #3.
 DE
 XX
 XX Alzheimer's disease; gamma-secretase; integral-membrane protein;
 KW beta-amyloid precursor protein; betaAPP; presenilin; PSI.
 KM
 XX
 XX Synthetic.
 OS
 XX
 XX Key Location/Qualifiers
 FH Modified-site 1
 FT /label="Nle
 FT /note="This residue is absent in the sequence
 FT shown in sequence listing"
 FT
 XX
 XX Modified-site 16
 FT /note="C-terminal amide"
 FT
 XX
 XX WO2001/5435-A2.
 PN
 XX
 XX 11-OCT-2001.
 PD
 XX
 XX 30-MAR-2001; 2001WO-US10453.
 PF
 XX
 XX 03-APR-2000; 2000US-194495P.
 PR
 XX
 XX (BRIM) BRISTOL-MYERS SQUIBB CO.
 PA
 XX
 XX Roberts SB, Hendrick JP, Vinitzky A, Lewis M, Smith DW, Pak R;
 PI WPI: 2001-648575/74.
 DR
 XX
 XX Novel gamma secretase protein, useful in the production of amyloids, is
 PT capable of cleaving beta-amyloid precursor protein to produce beta
 PT amyloid peptide -
 XX
 XX Example 8; Page 83; 127pp; English.
 PS
 XX
 XX The invention relates to the field of plaque amyloid deposits that are
 CC the hallmarks of Alzheimer's disease. In particular, the invention
 CC relates to an isolated, functionally-active protein that has

CC gamma-secretase activity. Gamma-secretase activity is necessary for
CC amyloid production. The present invention also relates to methods for
CC isolating integral-membrane proteins and protein complexes, including
CC the gamma-secretase protein of the invention. The method is useful for
CC monitoring the cleavage of beta-amyloid precursor protein (betaAPP)
CC by gamma-secretase. The present sequence is a synthetic peptide
CC antigen of presenilin peptide, PSL. This sequence is used in the
CC exemplification of the invention.
SQ Sequence 16 AA;
Query Match 86.7%; Score 13; DB 22; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.6e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHLGPHRSTPEER 13
DB 4 SHLGPHRSTPEER 16
RESULT 70
AAU58947
ID AU58947 standard; Protein; 85 AA.
XX
AC AAU58947;
XX
DT 13-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #19843.
XX
KM SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KM uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KM inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KM dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199947P.
PR 02-JUN-2000; 2000US-208641P.
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIXA CCRP.
XX
PI Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59596.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
PS Example 1; SEQ ID NO 20142; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies

CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 85 AA;
Query Match 46.7%; Score 7; DB 22; Length 85;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 LGPHRST 9
DB 26 LGPHRST 32
RESULT 71
AAG44010
ID AAG44010 standard; Protein; 93 AA.
XX
AC AAG44010;
XX
DT 18-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 55075.
XX
KM Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence; corn.
XX
OS Zea mays subsp. mays.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149426.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154409.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157753.
PR 05-OCT-1999; 99US-0157757.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158269.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160960.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.

PR 28-OCT-1999: 99US-0161992.
PR 28-OCT-1999: 99US-0161993.
PR 29-OCT-1999: 99US-0162142.

Query Match

Best Local Similarity 46.7%; Score 7; DB 21; Length 93;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRAA 15
Db 25 TPESRAA 31

RESULT 72

AAU41658
ID AAU41658 standard; Protein; 107 AA.

AC AAU41658;

DT 13-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #2554.

XX SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX dermatological; osteopathic; neuroprotectant.

OS Propionibacterium acnes.

PN WO200181581-A2.

PD 01-NOV-2001.

PF 20-APR-2001; 2001WO-US12865.

PR 21-APR-2000; 2000US-199047P.

PR 02-JUN-2000; 2000US-208841P.

PR 07-JUL-2000; 2000US-216747P.

PA (CORI-) CORNMA CORP.

PI Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;

PI L'malsoneuve J, Zhang Y, Jen S, Carter D;

DR WPI: 2001-616774/71.

DR N-PSDB; AAS59515.

XX Propionibacterium acnes polypeptides and nucleic acids useful for

PT vaccinating against and diagnosing infections, especially useful for

PT treating acne vulgaris.

XX Example 1; SEQ ID No 2853; 1069pp; English.

CC Sequences AAU93105-AAU6017 represent Propionibacterium acnes immunogenic

CC polypeptides. The proteins and their associated DNA sequences are used in

CC the treatment, prevention and diagnosis of medical conditions caused by

CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,

CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.

CC P. acnes is also involved in infections of bone, joints and the central

CC nervous system, however it is particularly involved in the inflammatory

CC lesions associated with acne vulgaris. A method for detecting the

CC presence or absence of P. acnes in a patient comprises contacting a

CC sample with a binding agent that binds to the proteins of the invention

CC and determining the amount of bound protein in the sample. The

CC polypeptides may be used as antigens in the production of antibodies

CC specific for P. acnes proteins. These antibodies can be used to

CC downregulate expression and activity of P. acnes polypeptides and

CC therefore treat P. acnes infections. The antibodies may also be used as

CC diagnostic agents for determining P. acnes presence, for example, by

CC enzyme linked immunosorbent assay (ELISA).

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 107 AA;

Query Match

Best Local Similarity 46.7%; Score 7; DB 22; Length 107;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPESR 13
Db 63 RSTPESR 69

RESULT 73

AAU04044
ID AAU04044 standard; Protein; 241 AA.

AC AAU04044;

DT 23-OCT-2001 (first entry)

DE Streptococcus coelicolor MmyT protein.

XX SCP1: methylenomycin cluster; mmc; MmyR; MmF; MmH;
XX MmF; MmR; MmyO; MmyG; MmyD; Mmr; heterologous gene expression..
XX Streptococcus coelicolor.

PN WO200148228-A1.

PD 05-JUL-2001.

PF 20-DEC-2000; 2000WO-GB04972.

PR 23-DEC-1999; 99GB-0030477.

PA (PLAN-) PLANT BIOSCIENCE LTD.

PI Chater KF, Bruton CJ, O'Rourke SJ, Wietzorrek AW;

DR WPI: 2001-425675/45.

DR N-PSDB; AAS07627.

DR AAU04045, AAU04046.

XX Novel expression cassette for expressing a nucleic acid of interest,

XX derived from the regulatory region of methylenomycin gene cluster of

XX SCP1 plasmid of Streptomyces coelicolor A3(2).

XX Claim 25; Fig 8f; 142pp; English.

XX The sequence represents the MmyT protein encoded by the mmyT gene carried

XX on the expression cassette present on plasmid SCP1. The expression

XX cassette is the regulatory region of the methylenomycin cluster (mmc)

XX from Streptomyces coelicolor A3(2), which encodes the MmyR, MmF, MmH,

XX MmF, MmR, MmyO, MmyG, MmyD and partial Mmr polypeptides. The

XX expression cassette is useful for expressing a nucleic acid of interest,

XX substantially only when the host cell culture reaches high cell density

XX at or close to the stationary phase of host cell culture. In particular

XX the system is useful in regulating methylenomycin production. Reduced or

XX no expression of the nucleic acid of interest is observed earlier in

XX growth, avoiding toxic effects of some gene products on growth and the

XX system does not require addition of exogenous inducer. The methylenomycin

XX cluster naturally present on a highly transmissible plasmid permits

XX properly regulated expression in diverse streptomycetes host and the

XX expression is driven by a strong promoter, leading to high yield of the

XX desired end product.

XX Sequence 241 AA;

XX Query Match

XX Best Local Similarity 46.7%; Score 7; DB 22; Length 241;

XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 TPESRAA 15
|||||
Db 206 TPESRAA 212

RESULT 74

ID AAM97977
AC AAM97977 standard; Peptide; 12 AA.

XX AAM97977;
XX

DT 24-JAN-2002 (first entry)

DE Human peptide #1252 encoded by a SNP oligonucleotide.

XX Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
XX neuroprotective; antimicrobial; gene therapy; vaccine; amylose; cancer;
KW amyloid protein; angiotensin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease.

OS Homo sapiens.

PN WO200147944-A2.

PD 05-JUL-2001.

PF 28-DEC-2000; 2000WO-US35498.

PR 28-DEC-1999; 99US-0173419.

PR 27-DEC-2000; 2000US-0173419.

PA (CURA-) CURAGEN CORP.

PI Shinkets RA, Leach M;

DR WPI; 2001-465210/50.

PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -

PS Disclosure; Page 3942; 4143pp; English.

XX The present invention relates to oligonucleotides (see AAL26793-AAL34659)
CC encoding polymorphic variants of proteins related to amylases, amyloid
CC proteins, angiotensin, apoptosis related proteins, cadherin, cyclin,
CC polymerase, oncogenes, histones, kinases, colony stimulating factors,
CC complement related proteins, cytochromes, kinesins, cytokines,
CC interferons, interleukins, G-protein coupled receptors and thioesterases.
CC The present sequence is a peptide encoded by one such oligonucleotide.
CC The oligonucleotides and the peptides encoded by them may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate expression of the proteins listed above. Disorders that may
CC be prevented, diagnosed and/or treated include multifactorial diseases
CC with a genetic component, such as autoimmune diseases (e.g. rheumatoid
CC arthritis, multiple sclerosis, diabetes, systemic lupus erythematosus
CC and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,
CC brain, breast, colon and kidney, leukemia), diseases of the nervous
CC system and an infection of pathogenic organisms.

SO Sequence 12 AA;

Query Match 40.0%; Score 6; DB 22; Length 12;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHRS 8
|||||
Db 5 LGPHRS 10

RESULT 75

ID AAM97978
AC AAM97978 standard; Peptide; 13 AA.

XX AAM97978;
XX

DT 24-JAN-2002 (first entry)

DE Human peptide #1253 encoded by a SNP oligonucleotide.

XX Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
XX neuroprotective; antimicrobial; gene therapy; vaccine; amylose; cancer;
KW amyloid protein; angiotensin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease.

OS Homo sapiens.

PN WO200147944-A2.

PD 05-JUL-2001.

PF 28-DEC-2000; 2000WO-US35498.

PR 28-DEC-1999; 99US-0173419.

PR 27-DEC-2000; 2000US-0173419.

PA (CURA-) CURAGEN CORP.

PI Shinkets RA, Leach M;

DR WPI; 2001-465210/50.

PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -

PS Disclosure; Page 3942; 4143pp; English.

XX The present invention relates to oligonucleotides (see AAL26793-AAL34659)
CC encoding polymorphic variants of proteins related to amylases, amyloid
CC proteins, angiotensin, apoptosis related proteins, cadherin, cyclin,
CC polymerase, oncogenes, histones, kinases, colony stimulating factors,
CC complement related proteins, cytochromes, kinesins, cytokines,
CC interferons, interleukins, G-protein coupled receptors and thioesterases.
CC The present sequence is a peptide encoded by one such oligonucleotide.
CC The oligonucleotides and the peptides encoded by them may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate expression of the proteins listed above. Disorders that may
CC be prevented, diagnosed and/or treated include multifactorial diseases
CC with a genetic component, such as autoimmune diseases (e.g. rheumatoid
CC arthritis, multiple sclerosis, diabetes, systemic lupus erythematosus
CC and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,
CC brain, breast, colon and kidney, leukemia), diseases of the nervous
CC system and an infection of pathogenic organisms.

SO Sequence 13 AA;

Query Match 40.0%; Score 6; DB 22; Length 13;
Best Local Similarity 100.0%; Pred. No. 6.3;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHRS 8
|||||
Db 6 LGPHRS 11

RESULT 76

```

ABB38298
ID   ABB38298 standard; Peptide: 55 AA.
XX
AC   ABB38298;
XX
DT   04-FEB-2002 (first entry)
XX
DE   Peptide #5804 encoded by human foetal liver single exon probe.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00669.
XX
PR   04-FEB-2000; 2000US-0180312.
XX   26-MAY-2000; 2000US-0207456.
XX   30-JUN-2000; 2000US-0608408.
XX   03-AUG-2000; 2000US-0632366.
XX   21-SEP-2000; 2000US-0234687.
XX   27-SEP-2000; 2000US-0236359.
XX   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI: 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
XX   analyzing gene expression in human fetal liver -
XX
PS   Claim 27; SEQ ID NO 30933; 639pp + sequence listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for
XX   measuring human gene expression in a sample derived from human foetal
XX   liver. The single exon nucleic acid probes may be used for predicting,
XX   measuring and displaying gene expression in samples derived from human
XX   fetal liver. The present sequence is a peptide encoded by a single exon
XX   nucleic acid probe of the invention.
XX
CC   Note: The sequence data for this patent did not form part of the
XX   printed specification, but was obtained in electronic format directly
XX   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 55 AA;
XX
Query Match 40.0%; Score 6; DB 22; Length 55;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 3 LGPHRS 8
Db 10 LGPHRS 15

```

```

OS   Homo sapiens.
XX
PN   WO200157274-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00666.
XX
PR   04-FEB-2000; 2000US-0180312.
XX   26-MAY-2000; 2000US-0207456.
XX   30-JUN-2000; 2000US-0608408.
XX   03-AUG-2000; 2000US-0632366.
XX   21-SEP-2000; 2000US-0234687.
XX   27-SEP-2000; 2000US-0236359.
XX   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI: 2001-488899/53.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
XX   hearts -
XX
PS   Claim 15; SEQ ID NO 25249; 530pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes for
XX   measuring human gene expression in a sample derived from human heart (see
XX   ABA21535-ABA41305). The present sequence is a protein encoded by one such
XX   probe. The probes may be used for predicting, measuring and displaying
XX   gene expression in samples derived from the human heart via microarrays.
XX   By measuring gene expression, the probes are useful for predicting,
XX   diagnosing, grading, staging, monitoring and prognosing diseases of the
XX   human heart and vascular system e.g. cardiovascular disease,
XX   hypertension, cardiac arrhythmias and congenital heart disease.
XX
CC   Note: The sequence data for this patent did not form part of the printed
XX   specification, but was obtained in electronic format directly from WIPO
XX   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 55 AA;
XX
Query Match 40.0%; Score 6; DB 22; Length 55;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 3 LGPHRS 8
Db 10 LGPHRS 15

```

```

RESULT 77
ABB23479
ID   ABB23479 standard; Protein: 55 AA.
XX
AC   ABB23479;
XX
DT   23-JAN-2002 (first entry)
XX
DE   Protein #5478 encoded by probe for measuring heart cell gene expression.
XX
KW   Human; gene expression; heart; microarray; vascular system;
XX   cardiovascular disease; hypertension; cardiac arrhythmia;
XX   congenital heart disease.
XX

```


PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI
 DR WPI; 2001-483446/52.
 XX
 PT Single exon nucleic acid probes for analyzing gene expression in human
 PT brains -
 XX
 PS Example 4; SEQ ID NO: 31025; 650bp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention.
 CC
 SO Sequence 55 AA;

Query Match 40.0%; Score 6; DB 22; Length 55;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHRS 8
 | | | | |
 DB 10 LGPHRS 15

RESULT 79
 ID AAM71442
 ID AAM71442 standard; Protein; 55 AA.
 AC
 XX AAM71442;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human bone marrow expressed probe encoded protein SEQ ID NO: 31748.
 XX
 XX Human: bone marrow expressed exon; gene expression analysis; probe;
 KM microarray; cancer; leukemia; lymphoma; myeloma.
 XX
 OS Homo sapiens.
 XX
 PN WO200157276-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00668.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI
 DR WPI; 2001-488900/53.
 XX

PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human bone marrow -
 XX
 PS Example 4; SEQ ID NO: 31748; 658bp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukemia and myeloma. The present sequence is a
 CC protein encoded by one of the probes of the invention.
 CC
 SO Sequence 55 AA;

Query Match 40.0%; Score 6; DB 22; Length 55;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHRS 8
 | | | | |
 DB 10 LGPHRS 15

RESULT 80
 ID AAM19095
 ID AAM19095 standard; Protein; 55 AA.
 AC
 XX AAM19095;
 XX
 DT 12-OCT-2001 (first entry)
 XX
 DE Peptide #5529 encoded by probe for measuring cervical gene expression.
 XX
 XX Probe: human: microarray; gene expression; cervical epithelial cell;
 KM cervical cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO200157278-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00670.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI
 DR WPI; 2001-488901/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human cervical epithelial cells -
 XX
 PS Claim 27; SEQ ID NO 23921; 487bp; English.
 XX
 CC The present invention relates to human single exon nucleic acid probes
 CC (SEN; see AAI10068-AA128459). The present sequence is a peptide encoded
 CC by one such probe. The SENs are derived from human HeLa cells. The SENs
 CC can be used to produce a single exon microarray, which can be used for
 CC measuring human gene expression in a sample derived from human cervical
 CC epithelial cells. By measuring gene expression, the probes are therefore
 CC useful in grading and/or staging of diseases of the cervix, notably
 CC cervical cancer.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 55 AA;
 Query Match 40.0%; Score 6; DB 22; Length 55;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 3 LGPHRS 8
 Db 10 LGPHRS 15
 RESULT 81
 ID AAM31734
 AC AAM31734 standard; Protein: 55 AA.
 XX
 AC AAM31734;
 XX
 DT 17-OCT-2001 (first entry)
 XX
 DE Peptide #5771 encoded by probe for measuring placental gene expression.
 XX
 KW Probe: microarray; human; placenta; antenatal diagnosis;
 KW genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200157272-A2.
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00663.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-48897/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human placenta -
 XX
 PS Claim 27; SEQ ID No 32003; 654pp; English.
 XX
 CC The present invention relates to single exon nucleic acid probes (SENP;
 CC see AAI13135-AI157546). The present sequence is a peptide encoded by one
 CC such probe. The probes are useful for producing a microarray for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from human placenta. The probes are useful for antenatal diagnosis of
 CC human genetic disorders.
 XX
 SQ Sequence 55 AA;
 Query Match 40.0%; Score 6; DB 22; Length 55;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 3 LGPHRS 8
 Db 10 LGPHRS 15
 RESULT 82
 ABG41248

ID ABG41248 standard; Peptide: 55 AA.
 XX
 AC ABG41248;
 XX
 DT 19-AUG-2002 (first entry)
 XX
 DE Human peptide encoded by genome-derived single exon probe SEQ ID 30913.
 XX
 KW Human; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200186003-A2.
 PD 15-NOV-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00665.
 XX
 PR 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2002-114183/15.
 XX
 PT Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 XX
 PS Claim 27; SEQ ID No 30913; 634pp; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridize at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray, assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease

CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
CC pulmonary alveolar proteinosis, Kartagener syndrome, fibrocystic
CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
CC and hyaline membrane disease. The present sequence is a peptide/protein
CC encoded by a single exon probe of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 55 AA;
Query Match 40.0%; Score 6; DB 23; Length 55;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 3 LGPHRS 8
|||||
DB 10 LGPHRS 15
RESULT 83
ABG06203
ID ABG06203 standard; Protein; 77 AA.
XX
AC ABG06203;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #6194.
XX
KM Human: chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS70390.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
PS Claim 20: SEQ ID No 36562; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating

CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 77 AA;
Query Match 40.0%; Score 6; DB 22; Length 77;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 3 LGPHRS 8
|||||
DB 63 LGPHRS 68
RESULT 84
ABG06427
ID ABG06427 standard; Protein; 77 AA.
XX
AC ABG06427;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #6418.
XX
KM Human: chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS70614.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
PS Claim 20: SEQ ID No 36786; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX

SO Sequence 77 AA:

Query Match 40.0%; Score 6; DB 22; Length 77;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHRS 8
 Db 63 LGPHRS 68

RESULT 85
 AAU42555

ID AAU42555 standard; Protein: 80 AA.

XX AAU42555;

DT 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #3451.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KM uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KM dermatological; osteopathic; neuroprotectant.

OS Propionibacterium acnes.

XX WO200181581-A2.

PN 01-NOV-2001.

PD 20-APR-2001; 2001WO-US12865.

XX 21-APR-2000; 2000US-199047P.

PR 02-JUN-2000; 2000US-208841P.

PR 07-JUL-2000; 2000US-216747P.

XX (CORI-) CORIXA CORP.

PI Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhatla A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX WPI: 2001-616774/71.
 DR N-PSDB; AAS59518.

PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -

XX Example 1; SEQ ID No 3750; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies

CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX

SO Sequence 80 AA:

Query Match 40.0%; Score 6; DB 22; Length 80;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HLGPHR 7
 Db 14 HLGPHR 19

RESULT 86
 ABG14485

ID ABG14485 standard; Protein: 80 AA.

XX ABG14485;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #14476.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KM food supplement; medical imaging; diagnostic; genetic disorder.
 KW Homo sapiens.

OS WO200175067-A2.

PN 11-OCT-2001.

PD 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

XX WPI: 2001-639362/73.
 DR N-PSDB; AAS78672.

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

PS Claim 20; SEQ ID No 4484; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.

XX
SQ Sequence 80 AA;

Query Match 40.0%; Score 6; DB 22; Length 80;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHRS 8
Db 66 LGPHRS 71
|||||

RESULT 87
ABG21779

ID ABG21779 standard; Protein; 84 AA.

XX
AC ABG21779;

XX
DT 18-FEB-2002 (first entry)

XX
DE Novel human diagnostic protein #21770.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS85966.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

XX Claim 20; SEQ ID NO 52138; 103bp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human

CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.

XX
SQ Sequence 84 AA;

Query Match 40.0%; Score 6; DB 22; Length 84;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RSTPES 12
Db 47 RSTPES 52
|||||

RESULT 88
ABG06213

ID ABG06213 standard; Protein; 85 AA.

XX
AC ABG06213;

XX
DT 13-FEB-2002 (first entry)

XX
DE Novel human diagnostic protein #6204.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS70400.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

XX Claim 20; SEQ ID NO 36572; 103bp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC Note: The sequence data for this patent did not appear in the printed

CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 85 AA;
Query Match 40.0%; Score 6; DB 22; Length 85;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 3 LGPHRS 8
|||||
Db 47 LGPHRS 52
RESULT 89
ABG06199
ID ABG06199 standard; Protein; 89 AA.
XX
AC ABG06199;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #6190.
XX
DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN W0200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2801; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR N-PSDB; AAS70386.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
PS Claim 20; SEQ ID No 36558; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 89 AA;
Query Match 40.0%; Score 6; DB 22; Length 89;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 3 LGPHRS 8
|||||
Db 36 LGPHRS 41
RESULT 90
ABG14499
ID ABG14499 standard; Protein; 89 AA.
XX
AC ABG14499;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #14490.
XX
DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN W0200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR N-PSDB; AAS78686.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
PS Claim 20; SEQ ID No 44858; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 89 AA;

Query Match 40.0%; Score 6; DB 22; Length 89;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LGPHRS 8
| | | | |
Db 62 LGPHRS 67

RESULT 91

ABG26397
ID ABG26397 standard; Protein: 89 AA.

XX AC ABG26397;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #26388.

XX KM Human: chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI: 2001-639362/73.

XX DR N-PSDB; AAS90584.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity -

XX PS Claim 20; SEQ ID No 56756; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC (II). (II) is useful for generating antibodies against it, detecting or
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. ABG00010-ABG30377 represent novel human
XX CC diagnostic amino acid sequences of the invention.
XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.

XX SO Sequence 89 AA;

Query Match 40.0%; Score 6; DB 22; Length 89;

Best Local Similarity 100.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 TPESRA 14
| | | | |
Db 49 TPESRA 54

RESULT 92

ABG18825
ID ABG18825 standard; Protein: 93 AA.

XX AC ABG18825;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #18816.

XX KM Human: chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI: 2001-639362/73.

XX DR N-PSDB; AAS83012.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity -

XX PS Claim 20; SEQ ID No 49184; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC (II). (II) is useful for generating antibodies against it, detecting or
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. ABG00010-ABG30377 represent novel human
XX CC diagnostic amino acid sequences of the invention.
XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.

XX SO Sequence 93 AA;

Query Match 40.0%; Score 6; DB 22; Length 93;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHRS 8
|||||
Db 54 LGPHRS 59

RESULT 93
AAV40019
ID AAV40019 standard; Peptide; 96 AA.

AC AAV40019;

DT 18-NOV-1999 (first entry)

DE Peptide sequence derived from a human secreted protein.

XX Secreted protein; gene therapy; cancer; tumor; fetal deficiency;
KW neurodegenerative disorder; developmental abnormality; blood disorder;
KW immune system disease; autoimmune disease; leukemia; inflammation;
KW allergy; Alzheimer's disease; cognitive disorder; schizophrenia;
KW obesity; osteoporosis; arthritis; infection; AIDS; diabetes; asthma;
KW connective tissue disorder; transplant rejection; sepsis; acne;
KW psoriasis; cardiovascular disorder; reproductive disorder;
KW food additive; food preservative; storage capability.

XX Homo sapiens.

PN MO9943693-A1.

XX 02-SEP-1999.

PD 24-FEB-1999; 99MO-US03939.

PF 26-FEB-1998; 98US-0076051.

PR 26-FEB-1998; 98US-0076052.

PR 26-FEB-1998; 98US-0076053.

PR 26-FEB-1998; 98US-0076054.

PR 26-FEB-1998; 98US-0076057.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Olsen HS, Florence K, Brewer LA, Ebner R, Ruben SM, Rosen CA;

PI Duan RD;

XX WPI: 1999-550857/46.

PS Disclosure: Page 25; 246pp; English.

CC AAV4001-92 are derived from human secreted proteins. The
CC polynucleotides and their corresponding secreted polypeptides are useful
CC for preventing, treating or ameliorating medical conditions, e.g. by
CC protein or gene therapy. Pathological conditions can also be diagnosed by
CC determining the amount of the new polypeptides in a sample or by
CC determining the presence of mutations in the polynucleotide. Specific
CC uses include developing products for the diagnosis or treatment of
CC cancer, tumors, neurodegenerative disorders, developmental abnormalities
CC and fetal deficiencies, blood disorders, sepsis, diseases of the immune
CC system, autoimmune diseases, inflammation, allergies, Alzheimer's and
CC cognitive disorders, schizophrenia, obesity, osteoporosis, arthritis,
CC infections, AIDS, connective tissue disorders, transplant rejection,
CC diabetes, asthma, sepsis, acne, psoriasis, cardiovascular disorders,
CC and reproductive disorders. The polypeptides or polynucleotides can
CC also be used as food additives or preservatives, such as to increase
CC or decrease storage capabilities, fat content, lipid, protein,
CC carbohydrate, vitamins, minerals, cofactors or other nutritional
CC components.

XX Sequence 96 AA;

Query Match 40.0%; Score 6; DB 20; Length 96;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHLGP 6
|||||
Db 42 SHLGP 47

RESULT 94
ABP35429
ID ABP35429 standard; Protein; 102 AA.

AC ABP35429;

DT 08-JUL-2002 (first entry)

DE Human synthase-like ORF4402 protein, SEQ ID NO:8804.

XX Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
KW disease monitoring; cytokine; cell proliferation; cell differentiation;
KW immune modulation; haematopoiesis regulation; tissue growth;
KW angiogenesis; activin; inhibin; chemotactic; chemokine; haemostatic;
KW thrombolytic; tumour inhibition; bodily characteristics; fertility;
KW behaviour; cancer; proliferative disorder; neurological disorder;
KW cardiovascular disease; immune system disorder; organ transplantation;
KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
KW hypothyroidism; cholesterol ester storage disease; infection; vulvare;
KW vasotropic; antipsoriatic; antidiabetic; cyostatic; nootropic;
KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
KW cardiant; hypotensive; antihypertoid; antiinflammatory; immunomodulator;
KW dermatological; analgesic; virucide; antibacterial; fungicide.

XX Homo sapiens.

OS WO200190466-A2.

PN 29-NOV-2001.

PD 24-MAY-2001; 2001MO-US17076.

PF 24-MAY-2000; 2000US-206690P.

PR (CURA-) CURAGEN CORP.

PR Leach MD, Shinkets RA;

PR WPI: 2002-106200/14.

PR N-PSDB: ABN79455.

XX Novel human polypeptides and polynucleotides useful for diagnosing,
XX preventing and treating cardiovascular disease, neurodegenerative,
XX hyperproliferative disorders and disorders related to organ
XX transplantation

PS Claim 10; Page 2444; 2508pp; English.

XX Sequences ABP31028-ABP35561 represent 4534 novel human proteins
XX designated ORF (open reading frame) 1-4534, and sequences ABN75054-
XX ABN79587 represent cDNAs encoding them. The invention also encompasses
XX polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
XX referred to as ORFX) proteins, polynucleotides at least 85% identical to
XX the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
XX polynucleotides, the recombinant production of ORFX proteins, antibodies
XX specific for ORFX proteins, methods of detecting ORFX polynucleotides and
XX polypeptides, methods of screening for modulators of ORFX expression or
XX activity, and methods of screening individuals for a predisposition to an
XX ORFX-associated disorder. The ORFX proteins of the invention have a wide
XX range of biological activities, such as cytokine, cell proliferation,
XX cell differentiation, immune modulation, haematopoiesis regulation,
XX tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
XX chemokine activity, haemostatic activity, thrombolytic activity,
XX receptor/ligand, antiinflammatory activity, tumour inhibition activity,

CC and anti-infective activity, and may also be involved in the determination
CC of bodily characteristics, fertility and behaviour. ORFX proteins,
CC nucleic acids and antibodies may be used in the treatment of cancers,
CC other proliferative disorders such as psoriasis and benign tumours,
CC neurological disorders such as epilepsy and Alzheimer's disease,
CC cardiovascular diseases, immune system disorders, disorders related to
CC organ transplantation, disorders of tissue growth and regeneration,
CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
CC storage disease, and infectious diseases caused by viral, bacterial,
CC fungal and other pathogens. ORFX nucleic acids may also be used as a
CC source of primers and probes, in the detection of ORFX genomic sequences
CC or transcripts, in the identification and cloning of homologous
CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
CC nucleic acids may additionally be used to produce transgenic animals
CC which may be useful for studying the function and/or activity of ORFX
CC protein, and in drug screening. The ORFX proteins may also be used as
CC immunogens to generate specific antibodies, which are useful in the
CC diagnosis, treatment and monitoring of ORFX-associated diseases.

SQ Sequence 102 AA;

Query Match 40.0%; Score 6; DB 23; Length 102;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RSTPES 12
| | | | |
DB 2 RSTPES 7

RESULT 95

AAU39273

ID AAU39273 standard; Protein; 103 AA.

AC AAU39273;

DT 13-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #169.

KW SAPHO syndrome; synovitis; acne; pustulosis; hypertostis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.

OS Propionibacterium acnes.

PN WO200181581-A2.

PD 01-NOV-2001.

PF 20-APR-2001; 2001WO-US12865.

PR 21-APR-2000; 2000US-199047P.

PR 02-JUN-2000; 2000US-208841P.

PR 07-JUL-2000; 2000US-216747P.

PA (CORI-) CORIXA CORP.

PI Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

DR N-PSDB; AAS59506.

PT Propionibacterium acnes polypeptides and nucleic acids useful for

PT vaccinating against and diagnosing infections, especially useful for

PT treating acne vulgaris -

PS Example 1; SEQ ID NO 468; 1069pp; English.

CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic

CC polypeptides. The proteins and their associated DNA sequences are used in

CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertostis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 103 AA;

Query Match 40.0%; Score 6; DB 22; Length 103;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 STPEPR 13
| | | | |
DB 70 STPEPR 75

RESULT 96

ABG11452

ID ABG11452 standard; Protein; 105 AA.

AC ABG11452;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #11443.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

PI WPI: 2001-639362/73.

DR N-PSDB; AAS75639.

PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity -

PS Claim 20; SEQ ID NO 41811; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The

CC polynucleotides are also used in diagnostics as expressed sequence tags

CC	for identifying expressed genes. (I) is useful in gene therapy techniques
CC	to restore normal activity of (II) or to treat disease states involving
CC	(II). (II) is useful for generating antibodies against it, detecting or
CC	quantitating a polypeptide in tissue, as molecular weight markers and as
CC	a food supplement. (II) and its binding partners are useful in medical
CC	imaging of sites expressing (II). (I) and (II) are useful for treating
CC	disorders involving aberrant protein expression or biological activity.
CC	The polypeptide and polynucleotide sequences have applications in
CC	diagnostics, forensics, gene mapping, identification of mutations
CC	responsible for genetic disorders or other traits to assess biodiversity
CC	and to produce other types of data and products dependent on DNA and
CC	amino acid sequences. ABG00010-ABG30377 represent novel human
CC	diagnostic amino acid sequences of the invention.
CC	Note: The sequence data for this patent did not appear in the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pct_sequences.
XX	
SQ	Sequence 105 AA;
Query Match	40.0%; Score 6; DB 22; Length 105;
Best Local Similarity	100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	3 LGPHRS 8
Db	36 LGPHRS 41
RESULT 97	
ABG11454	
ID	ABG11454 standard; Protein; 107 AA.
AC	ABG11454;
XX	
DT	18-FEB-2002 (first entry)
XX	
DE	Novel human diagnostic protein #11445.
XX	
KW	Human; chromosome mapping; gene mapping; gene therapy; forensic;
KM	food supplement; medical imaging; diagnostic; genetic disorder.
XX	
OS	Homo sapiens.
XX	
PN	M020001.75067-A2.
PD	11-OCT-2001.
XX	
PF	30-MAR-2001; 2001WO-US08631.
PR	31-MAR-2000; 2000US-0540217.
PR	23-AUG-2000; 2000US-0649167.
XX	
PA	(HYSE-) HYSEQ INC
XX	
PI	Dremanac RT, Liu C, Tang YT;
XX	
DR	WPI; 2001-639362/73.
DR	N-PSDB; AAS75641.
XX	
PT	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity -
PS	
Claim 20; SEQ ID NO 41813; 103pt; English.	
CC	The invention relates to isolated polynucleotide (I) and
CC	polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC	polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC	and gene mapping, and in recombinant production of (II). The
CC	polynucleotides are also used in diagnostics as expressed sequence tags
CC	for identifying expressed genes. (I) is useful in gene therapy techniques
CC	to restore normal activity of (II) or to treat disease states involving

CC	(II). (I) is useful for generating antibodies against it, detecting or
CC	quantitating a polypeptide in tissue, as molecular weight markers and as
CC	a food supplement. (II) and its binding partners are useful in medical
CC	imaging of sites expressing (II). (I) and (II) are useful for treating
CC	disorders involving aberrant protein expression or biological activity.
CC	The polypeptide and polynucleotide sequences have applications in
CC	diagnostics, forensics, gene mapping, identification of mutations
CC	responsible for genetic disorders or other traits to assess biodiversity
CC	and to produce other types of data and products dependent on DNA and
CC	amino acid sequences. ABG00010-ABG30377 represent novel human
CC	diagnostic amino acid sequences of the invention.
CC	Note: The sequence data for this patent did not appear in the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pct_sequences.
XX	
SQ	Sequence 107 AA;
Query Match	40.0%; Score 6; DB 22; Length 107;
Best Local Similarity	100.0%; Pred. No. 39;
Matches 6; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	3 LGPHRS 8 79 LGPHRS 84
Db	
RESULT 98	
ABG06208	
ID	ABG06208 standard; Protein; 112 AA.
XX	
AC	ABG06208:
XX	
DT	13-FEB-2002 (first entry)
XX	
DE	Novel human diagnostic protein #6199.
XX	
KW	Human: chromosome mapping; gene mapping; gene therapy; forensic;
KM	food supplement; medical imaging; diagnostic; genetic disorder.
XX	
OS	Homo sapiens.
XX	
PN	WO800175067-A2.
PD	11-OCT-2001.
XX	
PF	30-MAR-2801; 2001WO-US08631.
XX	
PR	31-MAR-2000; 2000US-0540217.
PR	23-AUG-2000; 2000US-0649167.
XX	
PA	(HYSE-) HYSEQ INC.
PI	Drmnac RT, Liu C, Tang YT;
XX	
DR	WPI: 2001-639362/73.
DR	N-PSDB: AAS70395.
XX	
PT	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity -
XX	
PS	Claim 20; SEQ ID NO 36567; 103pp; English.
XX	
CC	The invention relates to isolated polynucleotide (I) and
CC	polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC	polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC	gene mapping, and in recombinant production of (II). The
CC	polynucleotides are also used in diagnostics as expressed sequence tags
CC	for identifying expressed genes. (I) is useful in gene therapy techniques
CC	to restore normal activity of (II) or to treat disease states involving
CC	(II). (II) is useful for generating antibodies against it, detecting or
CC	quantitating a polypeptide in tissue, as molecular weight markers and as

CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

SQ Sequence 112 AA;

Query Match 40.0%; Score 6; DB 22; Length 112;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHRS 8
|||||
Db 55 LGPHRS 60

RESULT 99
ABG06438
ID ABG06438 standard; Protein; 112 AA.
XX
AC ABG06438;
XX

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #6429.

KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.

PN WO00175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR N-PSDB; AAS70625.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX

PS Claim 20; SEQ ID No 36797; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating

CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

SQ Sequence 112 AA;

Query Match 40.0%; Score 6; DB 22; Length 112;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHRS 8
|||||
Db 55 LGPHRS 60

RESULT 100
ABG14494
ID ABG14494 standard; Protein; 112 AA.
XX
AC ABG14494;
XX

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #14485.

KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.

PN WO00175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR N-PSDB; AAS78681.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX

PS Claim 20; SEQ ID No 44853; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX

SQ Sequence 112 AA;

Query Match 40.0%; Score 6; DB 22; Length 112;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHRS 8
|||||
Db 55 LGPHRS 60

RESULT 101
ABG28253
ID ABG28253 standard; Protein; 112 AA.
XX
AC ABG28253;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #28244.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR N-PSDB; AAS92440.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 58612; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX

SQ Sequence 112 AA;

Query Match 40.0%; Score 6; DB 22; Length 112;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHRS 8
|||||
Db 55 LGPHRS 60

RESULT 102
ABG24883
ID ABG24883 standard; Protein; 118 AA.
XX
AC ABG24883;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #24874.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR N-PSDB; AAS89070.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 55242; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC amino acid sequences. ABG00010-ABG30377 represent novel human

CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

Sequence 118 AA;

Query Match 40.0%; Score 6; DB 22; Length 118;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHRS 8
| | | | |
Db 65 LGPHRS 70

RESULT 103
ABG27021

ID ABG27021 standard; Protein; 120 AA.

AC ABG27021;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #27012.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN MO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR N-PSDB; AAS91208.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

PS Claim 20; SEQ ID No 57380; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed

CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

Sequence 120 AA;

Query Match 40.0%; Score 6; DB 22; Length 120;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHRS 8
| | | | |
Db 17 LGPHRS 22

RESULT 104
ABG14479

ID ABG14479 standard; Protein; 125 AA.

AC ABG14479;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #14470.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN MO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR N-PSDB; AAS78666.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

PS Claim 20; SEQ ID No 44838; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

```
XX Sequence 125 AA;
SQ Query Match 40.0%; Score 6; DB 22; Length 125;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LGPHRS 8
DB 55 LGPHRS 60

RESULT 105
AAU22757
ID AAU22757 standard; Protein; 126 AA.
XX AAU22757;
AC
XX
XX 18-DEC-2001 (first entry)
DE Human prostate cancer antigen, Seq ID No 276.
XX
XX Human prostate cancer antigen, Seq ID No 276.
KW reproductive system; chromosomal marker; forensic; urinary disorder;
KM chronic nephritis; blood-related disorder; thrombosis.
XX
OS Homo sapiens.
XX
XX WO200155316-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01328.
XX
XX 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
XX 24-FEB-2000; 2000US-0184664.
XX 02-MAR-2000; 2000US-0186350.
XX 16-MAR-2000; 2000US-0189874.
XX 17-MAR-2000; 2000US-0190076.
XX 18-APR-2000; 2000US-0198123.
XX 19-MAY-2000; 2000US-0205515.
XX 07-JUN-2000; 2000US-0209467.
XX 28-JUN-2000; 2000US-0214886.
XX 30-JUN-2000; 2000US-0215135.
XX 07-JUL-2000; 2000US-0216647.
XX 07-JUL-2000; 2000US-0216880.
XX 11-JUL-2000; 2000US-0217487.
XX 11-JUL-2000; 2000US-0217496.
XX 14-JUL-2000; 2000US-0218290.
XX 26-JUL-2000; 2000US-0220963.
XX 26-JUL-2000; 2000US-0220964.
XX 14-AUG-2000; 2000US-0224518.
XX 14-AUG-2000; 2000US-0224519.
XX 14-AUG-2000; 2000US-0225213.
XX 14-AUG-2000; 2000US-0225214.
XX 14-AUG-2000; 2000US-0225266.
XX 14-AUG-2000; 2000US-0225267.
XX 14-AUG-2000; 2000US-0225268.
XX 14-AUG-2000; 2000US-0225270.
XX 14-AUG-2000; 2000US-0225447.
XX 14-AUG-2000; 2000US-0225757.
XX 14-AUG-2000; 2000US-0225758.
XX 14-AUG-2000; 2000US-0225759.
XX 18-AUG-2000; 2000US-0226279.
XX 22-AUG-2000; 2000US-0226681.
XX 22-AUG-2000; 2000US-0226868.
XX 22-AUG-2000; 2000US-0227182.
XX 23-AUG-2000; 2000US-0227009.
XX 30-AUG-2000; 2000US-0228924.
XX 01-SEP-2000; 2000US-0229287.
XX 01-SEP-2000; 2000US-0229343.
XX 01-SEP-2000; 2000US-0229344.

PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 27-SEP-2000; 2000US-0235837.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
```

[illegible]

XX	Human; reproductive system related antigen; reproductive system disorder
KW	cancer; gene therapy.
OS	Homo sapiens.
XX	
PN	WO200155320-A2.
XX	
PD	02-AUG-2001.
XX	
PF	17-JAN-2001; 2001WO-USO1339.
XX	
PR	31-JAN-2000; 2000US-0179065.
PR	04-FEB-2000; 2000US-0180628.
PR	24-FEB-2000; 2000US-0184664.
PR	02-MAR-2000; 2000US-0186350.
PR	16-MAR-2000; 2000US-0188874.
PR	17-MAR-2000; 2000US-0190076.
PR	18-APR-2000; 2000US-0198123.
PR	19-MAY-2000; 2000US-0205515.
PR	07-JUN-2000; 2000US-0209467.
PR	28-JUN-2000; 2000US-0214886.
PR	30-JUN-2000; 2000US-0215135.
PR	07-JUL-2000; 2000US-0215647.
PR	07-JUL-2000; 2000US-0216880.
PR	11-JUL-2000; 2000US-0217487.
PR	11-JUL-2000; 2000US-0217496.
PR	14-JUL-2000; 2000US-0218290.
PR	26-JUL-2000; 2000US-0220963.
PR	26-JUL-2000; 2000US-0220964.
PR	14-AUG-2000; 2000US-0224518.
PR	14-AUG-2000; 2000US-0224519.
PR	14-AUG-2000; 2000US-0225213.
PR	14-AUG-2000; 2000US-0225214.
PR	14-AUG-2000; 2000US-0225266.
PR	14-AUG-2000; 2000US-0225268.
PR	14-AUG-2000; 2000US-0225270.
PR	14-AUG-2000; 2000US-0225447.
PR	14-AUG-2000; 2000US-0225757.
PR	14-AUG-2000; 2000US-0225758.
PR	14-AUG-2000; 2000US-0225759.
PR	14-AUG-2000; 2000US-0226279.
PR	18-AUG-2000; 2000US-0226681.
PR	22-AUG-2000; 2000US-0226868.
PR	22-AUG-2000; 2000US-0227182.
PR	23-AUG-2000; 2000US-0227009.
PR	30-AUG-2000; 2000US-0228924.
PR	01-SEP-2000; 2000US-0229287.
PR	01-SEP-2000; 2000US-0229343.
PR	01-SEP-2000; 2000US-0229344.
PR	01-SEP-2000; 2000US-0229345.
PR	05-SEP-2000; 2000US-0229509.
PR	05-SEP-2000; 2000US-0229513.
PR	06-SEP-2000; 2000US-0230437.
PR	06-SEP-2000; 2000US-0230438.
PR	08-SEP-2000; 2000US-0231242.
PR	08-SEP-2000; 2000US-0231243.
PR	08-SEP-2000; 2000US-0231244.
PR	08-SEP-2000; 2000US-0231413.
PR	08-SEP-2000; 2000US-0231414.
PR	08-SEP-2000; 2000US-0232080.
PR	08-SEP-2000; 2000US-0232081.
PR	12-SEP-2000; 2000US-0231968.
PR	14-SEP-2000; 2000US-0232397.
PR	14-SEP-2000; 2000US-0232398.
PR	14-SEP-2000; 2000US-0232399.
PR	14-SEP-2000; 2000US-0232400.
PR	14-SEP-2000; 2000US-0232401.
PR	14-SEP-2000; 2000US-0233063.
PR	14-SEP-2000; 2000US-0233064.
PR	14-SEP-2000; 2000US-0233065.
PR	21-SEP-2000; 2000US-0234223.

9300.
160.
91.
0.
17-N
17-
PR
PR
PR
PR
PR

```

SQ Sequence 126 AA;
Query Match 40.0%; Score 6; DB 22; Length 126;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHLGPH 6
    |||||
Db 61 SHLGPH 66

```

AA001111

standard; protein; 134 AA.

AAC01111;

20 NOV 2001 (first entry)

..... FOR THE PURPOSES OF THIS ACT

XX

vaccine; peptide therapy; stem cell growth factors; cell differentiation; gene therapy;

... nervous system disorders; arthritis; inflammation.

XXXX

XXX

XX

38-FEB-2000

200005-03/1409

X

2

IN FOLD; MAILB1042.

diagnosing and treat-

... frequent drinking; English.

enclosed proteins (AA000010-AA013910) that exhibit activity of